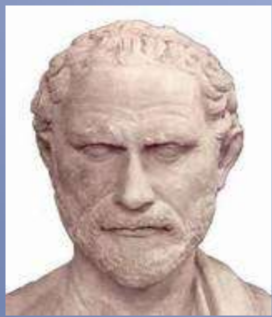


Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017

Elena Babatsouli, Editor

Institute of Monolingual and Bilingual Speech
Chania, Greece

PROCEEDINGS OF THE INTERNATIONAL SYMPOSIUM ON MONOLINGUAL AND BILINGUAL SPEECH 2017

EDITOR
ELENA BABATSOULI

INSTITUTE OF MONOLINGUAL AND BILINGUAL SPEECH
CHANIA, GREECE

Publisher: Institute of Monolingual and Bilingual Speech
Chania 73100, Greece

Symposium dates: 4-7 September 2017
Publication date: December 2017

ISBN: 978-618-82351-1-3
URL: http://www.ismbs.eu/publications-2017/

CONTENTS

ACKNOWLEDGMENTS

The editor wishes to express her gratitude to all the participants of the International Symposium on Monolingual and Bilingual Speech 2017 and its Proceedings: plenary speakers, invited speakers, contributed paper authors and speakers, members of the organizing committee, and members of the scientific committee. Special thanks to all for bringing valuable expertise to pave the way of research on monolingual and bilingual speech into the future.

FOREWORD

The Proceedings contain 50 papers from 150 presentations at the International Symposium on Monolingual and Bilingual Speech 2017 which took place in Chania, Greece on 4-7 September. The Symposium sprang from yearning for a specialized conference on speech that cuts across dividing boundaries between language subfields: first language, second language, bilingual, multilingual; child or adult; typical or impaired. The Symposium encourages investigations that go to the heart of matters, widening existing horizons and perspectives, kindling a holistic viewpoint, fostering collaborations across the board and, ultimately, sparking innovative thought and approaches. Participant affiliations covered 35 countries in Europe, North and South America, Africa, Asia, and Australia. Research included 45 languages: Afrikaans, Albanian, Arabic, Armenian, Bengali, Bosnian, Bulgarian, Cantonese, Catalan, Chinese, Czech, Danish, Dutch, English, Farsi, Finnish, French, Frisian, Georgian, German, Greek, Hausa, Hebrew, Hungarian, Inuktitut, isiXhosa, Italian, Japanese, Korean, Lithuanian, Mandarin, Miyaka Ryukyuan/Ikema, Norwegian, Nupe, Polish, Portuguese, Punjabi, Romanian, Russian, Spanish, Swedish, Turkish, Urdu, Vietnamese.

SCIENTIFIC COMMITTEE

Eric Raimy (University of Wisconsin, USA)
Lucrecia Rallo Fabra (Universitat de les Illes Balears, Spain)
Yvan Rose (University of Newfoundland, Canada)
Eirini Sanoudaki (Bangor University, UK)
James M. Scobbie (Queen Margaret University, UK)
Ellen Simon (University of Ghent, Belgium)
Anna Sosa (Arizona State University, USA)
Dimitrios Sotiropoulos (Technical University of Crete, Greece)
Joseph P. Stemberger (University of British Columbia, Canada)
Isao Ueda (Osaka University, Japan)
Virve-Anneli Vihman (University of Manchester, UK)
Magdalena Wrembel (Adam Mickiewicz University Poznań, Poland)
Mehmet Yavaş (Florida International University, USA)

Non-arbitrariness in Swedish parent-child interactions

Åsa Abelin
asa.abelin@gu.se

University of Gothenburg

Abstract. The purpose of this paper is to study the frequency and use of onomatopoeic and sound symbolic words, reduplications and emotional interjections in Swedish by analyzing the interactions between adults and small children in a longitudinal corpus. The study looks for correlations between the variables age at recording, type of non-arbitrary word, word frequency, and type of speech – IDS or child speech. The corpus used is the Lacerda corpus, which consists of recordings and transcriptions of three Swedish speaking children and their caregivers recorded at six occasions between approximately 1 and 2;8 years. Onomatopoeic, sound symbolic expressions, reduplications and emotional interjections in the speech of the children and the adults were analyzed. The analysis shows that more onomatopoeic words and reduplications are produced in the early ages than in the later ages of the children, and that the caregivers produce more non-arbitrary words and expressions than the children do. Most importantly it shows that there is an interaction between the adult and the child since the children's productions follow the adult's productions, at every recording session.

Keywords: onomatopoeia, sound symbolism, infant directed speech, interaction, longitudinal

Introduction

Non-arbitrary words can be onomatopoeic, sound symbolic or emotionally expressive. The type of non-arbitrary words and expressions which are considered in this study are the iconic onomatopoeic (e.g., *moo*, for the sound of a cow) and sound symbolic (e.g., *peep peep* for small objects and *hu hu* for big objects), as well as iconic reduplications (e.g., *jump jump*, imitating movement) and indexical emotional interjections (e.g., *oh*). By 'words and expressions', it is meant conventional non-arbitrary words and word forms which are more unconventional. It is not always easy to draw an exact line between conventional and unconventional words as conventionality is often implicitly defined with reference to written language. It should be pointed out that also non-arbitrary expressions are often conventionalized to a certain degree. The onomatopoeic and sound symbolic words are defined as words and expressions imitating non-speech sounds and speech sounds, and words and expressions imitating perceptual properties of other senses than hearing e.g., size, light, movement or form (Abelin, 1999). Non-arbitrariness can also be reflected in emotional prosody. Emotional interjections are considered indexical in the Peirceian sense, expressing psychological states of emotion or attitude. Non-arbitrary words and expressions are embodied in the sense that they are more directly grounded in sensory experience (hearing, touch, sight) than arbitrary words. These words are thus claimed to be easier to learn for the child – the sound symbolism boot strapping hypothesis (cf. Imai & Kita, 2014, as well as Maurer, Pathman, & Mondloch, 2006), but the reason for easy learnability could also be greater prosodic saliency (cf. Laing, Vihman & Keren-Portnoy, 2016). The purpose of the present paper is to study the frequency and use of onomatopoeic and sound symbolic words and expressions, reduplications and emotional interjections in child directed speech and in children's speech, in a longitudinal corpus of Swedish.

The research question is: What is the quantitative and qualitative development of onomatopoeic and sound symbolic expressions, reduplications and emotional interjections in Swedish children's speech and infant directed speech, and what is the role of interaction between child and adult?

Background

It is generally said that there is higher frequency of onomatopoeia and sound symbolic words in infant directed speech (IDS) than in adult speech directed to other adults. Imai and Kita (2014) claim that sounds symbolism and iconic representations, in general, help establish word-referent associations; word learning in 14-month-old Japanese speaking infants was facilitated by sound symbolism. They also present results that Japanese caretakers adjust their output according to the lexical level of the child, and that the caretakers use mimetics (a grammaticalized form of sound symbolism in Japanese) as a tool for the adjustment. Laing *et al*. (2016) compared onomatopoeic words acoustically with corresponding conventional words in infant directed speech. The issue was whether the greater prosodic saliency or the concept of the sound symbolism boot strapping hypothesis is what facilitates first language acquisition. They report that onomatopoeia produced more saliently than non-onomatopoeic words as regards pitch, duration, frequency (words produced more frequently due to reduplication) and word isolation. It is also generally accepted (cf. van Dijk, van Geert, Korecky-Kröll, Maillochon, Laaha, Dressler, & Bassano, 2013) that children develop their language in interaction with adults and that input is essential. The question is now what holds for non-arbitrary words and expressions in Swedish child-adult pairs.

Method

Material

The corpus used is the entire Lacerda corpus (Rose & MacWhinney, 2014), which consists of audio recordings and transcriptions of three Swedish speaking children and their caregivers recorded at six occasions between approximately 1 and 2;8 years of age at Stockholm University, in a naturalistic setting with the child interacting with one of its caregivers. The children were recorded in the same setting, with the same toys from which they could choose freely. In total, there are 18 recordings, between 13 and 57 minutes long. The recordings are orthographically transcribed and annotated, and the recordings are aligned with the transcriptions in a transcript browser.

The girl C was recorded from 1;0.17 to 2;7.14, the girl E was recorded from 1;1.8 to 2;8.27, and the girl T was recorded from 1;0.26 to 2;8.22. All ages of recordings for every child can be seen in figures 1, 2 and 4.

Method of analysis

The recordings were listened to by the author, in conjunction with reading of the transcriptions. All onomatopoeic and sound symbolic words and expressions, and reduplications were extracted, also those which had not been originally transcribed in the corpus. It is a fact that it is easy for transcribers to overlook non-conventional words. Emotional interjections were also extracted, but not those expressions where emotion was only expressed by prosody. The onomatopoeic and sound symbolic expressions, and emotional interjections in the speech of the children and in the child directed speech were categorized. These expressions were also categorized as reduplicated whenever that was the case. Reduplicated expressions can thus be onomatopoeic, sound symbolic or emotional interjections. The study analysed correlations between the variables: age at recording, type of non-arbitrary word, frequency of words and expressions, and type of speech (IDS or child speech). Prosody was not analysed.

Results

The types of non-arbitrary expressions found in the analysis were onomatopoeia and emotional interjections, all of them often reduplicated by both adults and children. There was a smaller amount of sound symbolic expressions – for the concepts size or movement, or as an intensifier. Sound symbolic words with [i:] for small things, and [u:] for large things occurred in the speech of both adults and children. Examples are [hu:hu:], [wu:], [pik pik], [pi:p pi:p]. Imitation of movement was not uncommon and was often expressed with reduplications of conventional words e.g., *hopp hopp* or *hoppi doppi* ‘jump jump’, klapp klapp klapp ‘stroke stroke stroke’, *kram kram* ‘hug hug’, *doing doing doing* (imitating movement). These were almost always produced by the adult.

In total, the adult produces more non-arbitrary expressions than the child, both quantitatively and qualitatively. This means that the adults produced more non-arbitrary tokens and that the adult produced a larger amount of different words than the child. This is true for every child-adult pair and for the three pairs taken together. Figure 1 shows that the adult produced more onomatopoeias at every recording session, and that the amount of child productions correlate with the adult productions.

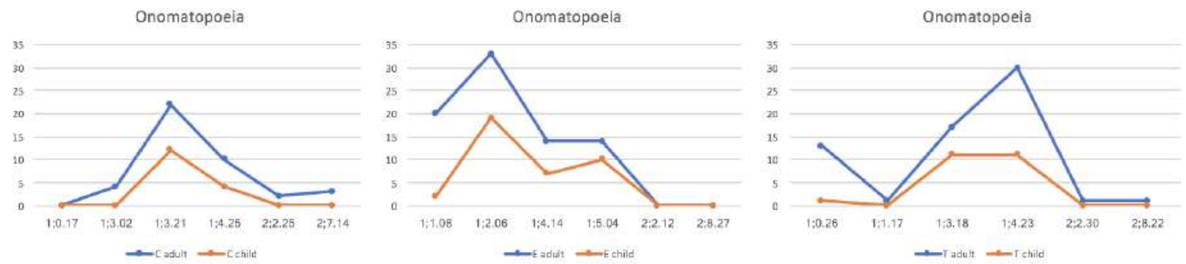


Figure 1. Number of onomatopoeic expressions (tokens) over six recording sessions for the three adult-child pairs

Examples of onomatopoeia are imitations of animal sounds e.g., *hoo hoo hoo* (owl), *mjau mjau* (cat), *voff voff* (dog) or *brum brum* (*vroom vroom* vehicles) but also of own sound mimicry e.g., *namnam* '*yummy*'. All correlations shown in Figure 1 are significant. R^2 for onomatopoeia C: 0.9704, E: 0.7139, T: 0.7685. We can see that the E-pair has a larger number of onomatopoeias at their maximum, compared to the C-pair and the T-pair. Figure 2 on reduplications shows a similar pattern as Figure 1, as concerns both correlations and the sessions with maximum number of productions.

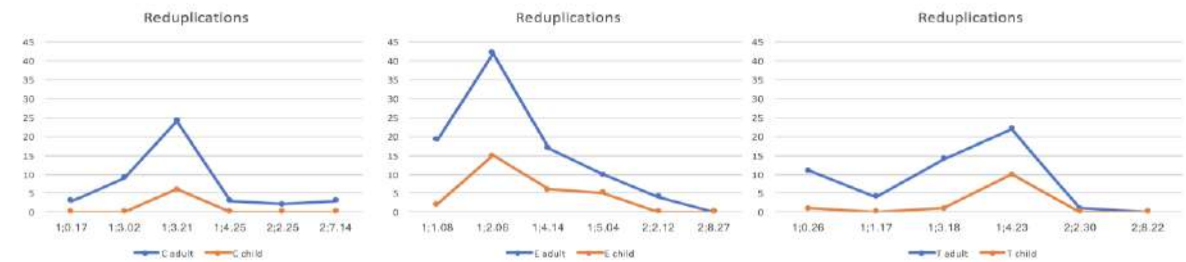


Figure 2. Number of reduplicated non-arbitrary expressions, i.e. onomatopoeia, sound symbolic expressions or emotional interjections, (tokens) over six recording sessions for the three adult-child pairs

All types of non-arbitrary expressions were reduplicated: onomatopoeic words and expressions, emotional interjections or imitations of movement, and the reduplications express the concepts of repetition, prolongation or intensification. All correlations shown in Figure 2 are significant. R^2 for reduplication C: 0.9124, E: 0.8557, T: 0.6939. We can again see that the E-pair has a larger number of reduplications at the maximum point, compared to the C-pair and the T-pair. Thus, there are significant correlations between the amounts of adult and child productions of onomatopoeia and reduplications at all ages. At each recording session, the child produces onomatopoeia and reduplicated expressions in relation to the quantity of the adults' productions. Therefore, it is not the case that the number of non-arbitrary expressions of the child is decreasing or increasing only with age. The children's ages for maximum number of non-arbitrary productions vary between the three pairs (see Figures 1 and 2). However, there is very little onomatopoeia and reduplications in the two or three last sessions. Relating maximum amount of onomatopoeia and reduplications with maximum number of turns for each of the 3x6 sessions, there is no correlation between number of turns and number of non-arbitrary productions. Thus, the number of onomatopoeia and reduplications is not an effect of a greater number of turns.

When we study the interaction between the adult and the child it is clear that the adult is generally introducing non-arbitrary expressions before the child's productions. This is quantified in Figure 3.

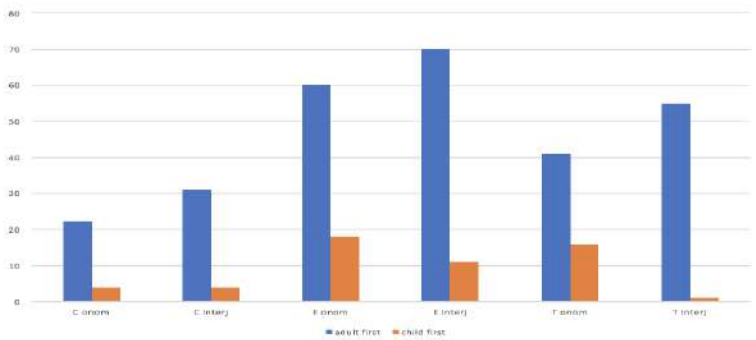


Figure 3. Number of first introductions of onomatopoeic words and emotional interjections for the three adult-child pairs

Figure 3 also shows that the difference between each adult and child is larger for the emotional interjections than for the onomatopoeic words and expressions. Each adult also produces more first emotional interjections than first onomatopoeias. As a contrast, children E and T each produce more first onomatopoeias than first emotional interjections but, all in all, the children produce much fewer onomatopoeias and emotional interjections in comparison to the adult.

In contrast to onomatopoeia and reduplications, the children began to produce emotional interjections late (see Figure 4). Another difference compared to the production of onomatopoeias and reduplications is that the productions of emotional interjections shown in Figure 4 do not exhibit any significant correlation between adults and children.

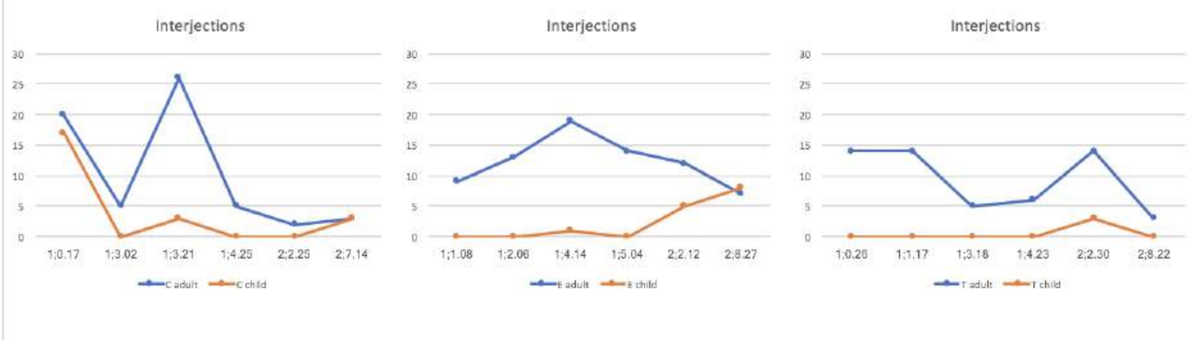


Figure 4. Number of emotional interjections (tokens) over six recordings sessions for the three adult-child pairs. Children produce emotional interjections mainly during one or two of the last two sessions

To be noted, the high number of emotional interjections at the first session for child C is due to a long interaction between the child and the adult scaring each other with the interjection *bu!* ‘booh!’. This is not really an interjection of emotion, but an interjection for evoking an emotion. Examples of common emotional interjections in the different recordings are: *oj* ‘oops’, *aj* ‘ouch’. In all recordings, there is great variation in the production of unconventional forms, primarily in the speech of the children, but also in adult speech e.g., imitations of horses, pigs, monkeys or cars. Cows were imitated by child C with e.g., [ɛmɛmɛmɛ], horses with [ø:øh], and dogs with [hø:hø], cars were imitated by child C with e.g., [vmmmmmm], [vmmm], [ɛummmm], [mmmmmmmmm] and by C’s father with e.g., [rrrawm] and [aauuun], all with very expressive prosody. E imitates the dog with [fhfhvufuffhøhi] and the monkey with [ihuhuhuhu]. These are all broad transcriptions. None of these onomatopoeic forms are conventional words in Swedish. The results thus show that more onomatopoeic words and expressions, and reduplications are produced to a larger extent in the early recordings than in the later recordings, and that the caregivers produce more onomatopoeic and sound symbolic words, reduplications and emotional interjections than the children do, at every recording session. The adult also in general produces the first occurrence of a word or expression. Impressionistic observations of the interactions indicate that the adult’s non-arbitrary expressions might sometimes work as attention getters when the child becomes bored of the situation, wants to take off the recording vest, walks around, falls over, hurts herself or starts moaning, etc. The parent naturally wants the child to continue the recording session, but the situation is not too different from similar settings in real life.

Discussion

It is evident from this study that the adults both produce more non-arbitrary expressions than the children, and more first introductions of these. The children's productions are following the corresponding productions of the adults. Since the adult is leading the interaction, the fewer onomatopoeic expressions in the later sessions can be interpreted as the adult calibrating and revising the language level of the child (cf. van Dijk *et al*., 2013). Fine-tuning of child directed speech and specifically the reduction in onomatopoeia with age has already been found in studies of Japanese children (Ogura, Murase & Yamashita, 1992; Ogura, Yoshimoto & Tsubota, 1997). What is certain from the present study is that the interaction with the adult seems to be crucial for the children's productions of non-arbitrary expressions. This is not different from children's word acquisition in general, but the children seem to be sensitive to the adult's production of onomatopoeia and reduplications at the earliest stages (cf. Laing, 2014), however not to emotional interjections. Why emotional interjections are not acquired as early as onomatopoeia and reduplications, it was an unexpected result. An explanation to this might lie in the fact that emotional interjections are not iconic but indexical (in the Peirceian sense), which could indicate that iconicity is a more fundamental non-arbitrary relation than indexicality. It can be noted here that many of these onomatopoeic and reduplicated word forms in Swedish are generally not used between adults otherwise than in playful or affectionate situations; however, there is need for more research on this question. It is not yet possible to decide between the sound-symbolic boot strapping hypothesis or the prosodic saliency hypothesis from this study, since analyses of prosody were not made. As prosody (intonation, prolongation, intensity) is very salient when one is listening to the recordings, this will be subjected to further research by analysing F0, durations and speech tempo; the extracted non-arbitrary expressions will be analysed prosodically and compared with arbitrary expressions of the same speakers. The role of gestures will also be addressed.

Conclusions

- The Swedish-speaking adults produce more onomatopoeic and sound symbolic words than the Swedish-speaking children do.
- Interaction with the Swedish-speaking adult triggers the Swedish-speaking children's productions of onomatopoeia and reduplications.
- Interaction means more for productions of onomatopoeia and reduplications in Swedish than the exact age of the child.
- Onomatopoeia is in part learned. Thus, onomatopoeia is not so different from other L1-acquisition constructs – the child is learning from the adult.
- However, iconicity can help the child connect the world to words, and therefore iconic expressions like onomatopoeia and reduplications occur early in language acquisition.
- The children's production of emotional interjections did not occur as early as their productions of onomatopoeia and reduplications, and they did not show alignment with the adults' productions.

References

Abelin, Å. (1999). *Studies in sound symbolism*. Göteborg: Gothenburg Monographs in Linguistics 17.

Dijk, M. van, Geert, P. van, Korecky-Kröll, K., Maillochon, I., Laaha, S., Dressler, W. U., & Bassano, D. (2013). Dynamic adaptation in child-adult language interaction. *Language Learning*, *63*, 243-270.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical transactions of the royal society*, *369*, 20130298.

Laing, C. E., (2014). A phonological analysis of onomatopoeia in early word production. *First Language*, *34*(5), 387-405.

Laing, C. E., Vihman, M., & Keren-Portnoy, T. (2016). How salient are onomatopoeia in the early input?: A prosodic analysis of infant-directed speech. *Journal of Child Language*, *1*(23).

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: sound–shape correspondences in toddlers and adults. *Developmental Science*, *9*(3), 316-322.

Ogura, T., Murase, T., & Yamashita, Y. (1992). Shoki gengo hattatsu ni kansuru chousa (1): Yoojigo kara seijingo e [A study of early language development (1): From Baby-word to Adult-word]. Shimane daigaku kyouikugakubu kiyoo [Buletin of the Faculty of Education, Shimane University], *26*, 57-63. http://ir.lib.shimane-u.ac.jp/ja/journal/E-MLS/26/--/article/1019

Ogura, T., Yoshimoto, Y., & Tsubota, M. (1997). Hahaoya no ikujigo to kodomo no gengo hattatsu ninchi hattatsu [Baby talk and children's linguistic and cognitive development]. Kobe daigaku hattatsugakubu kenkyuu kiyoo [Buletin of the Faculty of Human Development, Kobe University], *5*, 1-14. http://www.lib.kobe-u.ac.jp/infolib/meta_pub/G0000003kernel_81000264

Rose, Y., & MacWhinney, B. (2014). *The PhonBank project*, Lacerda corpus, PhonBank/TalkBank.org.

To cite this article: Abelin, Å. (2017). Non-arbitrariness in Swedish parent-child interactions. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 1-6). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

The acquisition of object relative clauses in Greek: The role of lexical NP restriction and featural mismatch

Christina Alexandri[1], Sofia Apostolopoulou[2], Admiljan Drobaniku[3], Eleni Kotroni[4], Spyridoula Varlokosta[5]
calexandri@phil.uoa.gr, sapost@phil.uoa.gr, admiljan@gmail.com, elkotroni@gmail.com, svarlokosta@phil.uoa.gr

National and Kapodistrian University of Athens

Abstract. Cross-linguistic studies have shown that children find object relatives harder than subject relatives and interpret this asymmetry in terms of Rizzi's (1990) principle of relativised minimality (RM) (e.g., Adani, van der Lely, Forgiarini, & Guasti, 2010; Belletti, Friedmann, Brunato, & Rizzi, 2012; Friedmann, Belletti, & Rizzi, 2009; Varlokosta, Nerantzini, & Papadopoulou, 2014, among others). Friedmann *et al*. (2009) propose that structural dissimilarity in terms of lexical NP restriction between the moved constituent and the intervening subject improves children's performance, while other studies (Adani *et al*., 2010; Belletti *et al*., 2012; Varlokosta *et al*., 2014) suggest that the feature specification of the moved and intervening element affects children's performance, as well. The present study investigates the effect of lexical NP restriction and featural mismatch on Greek-speaking children's comprehension of object relatives. Our results indicate that not only lexical NP restriction but also the featural mismatch between the moved and intervening DP in terms of number mismatch affect children's performance (Adani *et al*., 2010). Interestingly, children's performance was better in the number mismatch condition in which the intervening subject was an overt lexically restricted DP instead of an impersonal pro with arbitrary interpretation. In addition, the presence of a resumptive clitic in the original merge position of the relative clause did not improve but, in some cases, deteriorated performance in children and in the adult control group.

Keywords: object relatives, Relativized Minimality, lexical NP restriction, number mismatch, arbitrary pro intervener

Introduction

Object relative clauses (ORCs) have been shown to be harder for children to comprehend compared to subject relatives clauses (SRCs) (for typically developing children: Crain, McKee, & Emiliani, 1990; de Villiers, Tager Flusberg, Hakuta, & Cohen, 1979; Friedmann, Belletti, & Rizzi, 2009; Sheldon, 1974, among many others).

A number of studies interpret this selective difficulty with ORCs within the Relativised Minimality (hereafter, RM) approach to non-argument dependencies (A'-dependencies). RM predicts that in a configuration like (1), a structural local relation between two constituents X and Y cannot hold when Z, a third constituent with similar morphosyntactic features, intervenes and is somehow a potential candidate for the local relation (Rizzi, 1990, 2001).

(1) … X … Z … Y …

Friedmann *et al*. (2009) explain the difficulties that Hebrew-speaking children (age range: 3;7-5;0 years) face with object dependencies in different constructions involving A'-movement (relative clauses and *wh*-questions) within the RM approach. They propose that children cannot establish a distant dependency between an A'-target and its copy when the A'-target DP and the intervening subject DP are both lexically restricted (i.e. specified with a complex feature set such as [D NP]). According to the authors, the minimality/intervention effects observed in object dependencies of different A'-constructions are due to a much stricter version of RM that children apply compared to adults, which requires feature disjointedness between the A'-target and the intervener. Specifically, in child grammar a disjointed relation between the A'-target and the intervener (i.e. +A … +B … <+A>) instantiates more successfully a local relation between the A'-target and its copy, because it is less costly in terms of its computation compared to a subset relation (i.e. +A, +B … +A … <+A, +B>). Under these assumptions, Friedmann *et al*. (2009) explain the following asymmetries observed in the acquisition of object dependencies in Hebrew: better performance on bare *wh*-questions and free relatives, where the moved DP is not lexically restricted, compared to *which-NP* questions and headed relatives, where both the moved and the intervening DP are lexically restricted. Furthermore, the alleviating effect that a (non-lexical) arbitrary pro intervening subject had on children's comprehension of ORCs in Hebrew is also explained within the RM approach. Nevertheless, Friedmann *et al*. (2009) leave open the issue of whether it is the presence of an arbitrary pro or the absence of lexical NP restriction that facilitates children's performance in this case.

Subsequent studies have investigated whether it is exclusively the lexical restriction [+ NP], as a whole, that triggers intervention or if different features of the DP can contribute to intervention in different ways. Adani, van der Lely, Forgiarini, and Guasti (2010) showed that Italian-speaking children comprehend ORCs better with number than with gender mismatch between the head and the subject of the relative clause. Thus, they proposed that intervention effects are sensitive not only to the presence of lexical restriction in the DP but also to the internal structure of the DP and that 'external and syntactically active features, such as Number, reduce intervention, whereas internal and (possibly) lexicalized features, such as Gender, do so to a lesser extent' (Adani *et al*., 2010:2.148).

Belletti, Friedmann, Brunato, and Rizzi (2012) assessed the effect of gender morphology on children's comprehension of ORCs in Hebrew and Italian. Whereas the gender mismatch between the moved and the intervening DP significantly improved the comprehension of ORCs in Hebrew, it did not in Italian. The authors attributed these results to the different status of gender in the two languages: in Hebrew, tensed verbs are inflected for gender, so gender is part of the featural composition of the clausal inflectional head, thus it is part of the feature set attracting the subject; in Italian tensed verbs are not inflected for gender, so gender is not part of the feature set attracting the subject. On the basis of these results, the authors concluded that only features that function as attractors for syntactic movement will enter into the computation of intervention.

Varlokosta, Nerantzini, and Papadopoulou (2014) investigate the comprehension of subject and object dependencies in different *wh*-movement structures in Greek and argue that although intervention effects emerge when both the moved and intervening DPs are lexically restricted, such effects also depend on the feature specification of the two DP constituents.

The present study examined the effect of lexical NP restriction and featural mismatch in the comprehension of ORCs in child Greek. Along the lines of previous work on intervention effects in child language, we predicted that children's performance will be higher when (a) the moved and intervening DPs differ in terms of their lexical NP restriction (Friedmann *et al*., 2009; Varlokosta *et al*., 2014 for Greek); (b) the moved and intervening DPs differ in their number specification (Adani *et al*., 2010). Furthermore, the role of a resumptive clitic (a weak type of pronoun which restates the antecedent in the ORC) at the original merge position and the effect of an intervening arbitrary pro subject were also assessed.

Method

Participants

Thirty-eight (38) monolingual Greek-speaking typically developing children (20 boys and 17 girls), aged from 3;1 to 4;11 years (mean age: 4;0, SD: 0.6), with no diagnosed language, hearing or speech pathologies, took part in the experiment. Their mental and linguistic profiles were assessed via two baseline tasks (Raven's Coloured Progressive Matrices, 1998; for Greek: Sideridis, Antoniou, Simos, & Mouzaki, 2015; PPVT-R, Dunn & Dunn, 1981; for Greek: Simos, Sideridis, Protopapas, & Mouzaki, 2011). Children were recruited from three private kindergartens in Athens.

Thirty-seven (37) monolingual Greek-speaking adults (17 men, 20 women), aged from 20 to 38 years (mean age: 25;3, SD: 3.6), were used as a control group to ensure that the experimental sentences were properly comprehended. No history of speech and language difficulties or any psychological, social or emotional disturbance was reported.

Materials

Comprehension of ORCs was assessed via an off-line Picture Selection (PS) task, which included six experimental conditions and one control condition (see Table 1). In *C1* and *C2*, both the moved and the intervening DPs contained lexical NP restriction, while in *C3* and *C4*, the moved DP did not include a lexical NP restrictor. *C2* and *C4* contained a resumptive object clitic, as well, in the original merge position. In order to investigate the role of featural mismatch in enhancing children's comprehension of ORCs, *C5* contained a number mismatch between the moved DP (singular) and the intervening subject DP (plural), while *C6* contained an arbitrary pro subject intervener and a resumptive object clitic in the original merge position. *C7* was a control condition with canonical argument order.

The task included 42 sentences with the verbs *pull*, *wave*, *photograph*, *chase*, *look*, and *bite* (six experimental sentences/verbs per condition). For the first three verbs, human characters were used (man, doctor, sailor, grandpa, policeman, and photographer) and for the second three, non-human characters (dog, rabbit, wolf, squirrel, crocodile, and hippo). The two DPs in each sentence were chosen on the basis of their written frequency in a Greek corpus (http://hnc.ilsp.gr/) and, in order to avoid one character being more prominent than the others, they had to meet the following criteria: they started with different sounds, they had the same number of syllables and similar frequencies in the corpus. Their pragmatic plausibility in the sentence was also taken into account. For example, in the sentence *o scilos ciniγa to laγo* 'the dog_NOM is chasing the rabbit_ACC', both DPs consist of two syllables, have similar frequency in the corpus, start with different sounds, and the event described in the sentence is a pragmatically plausible one. The experimental sentences were chosen on the basis of a familiarity test administrated to 98 monolingual Greek-speaking adults (30 men, 68 women, mean age: 27;4, SD: 7.36).

Table 1. Experimental conditions

A/a	Conditions	Examples
C1	+NP restriction	ðikse mu **to laγo$_i$** pu ciniγa **o scilos** t_i 'show me the rabbit_ACC that is-chasing the dog_SNOM'
C2	+NP restriction, + clitic	ðikse mu **to laγo$_i$** pu **ton$_i$** ciniγa **o scilos** t_i 'show me the rabbit_ACC that CL_ACC is-chasing the dog_NOM'
C3	-NP restriction	ðikse mu **afton$_i$** pu ciniγa **o scilos** t_i 'show me whom that is-chasing the dog_NOM'
C4	-NP restriction, +clitic	ðikse mu **afton$_i$** pu **ton$_i$** ciniγa **o scilos** t_i 'show me that one CL_ACC is-chasing the dog_NOM'
C5	number mismatch, +NP restriction	ðikse mu **to laγo$_i$** pu ciniγun **i scili** t_i 'show me the rabbit_ACC that are-chasing the dogs_NOM'
C6	number mismatch, +pro, +clitic	ðikse mu **to laγo$_i$** pu **ton$_i$** ciniγun **pro$_j$** t_i 'show me the rabbit_ACC that are-chasing'
C7	control	o scilos ciniγa to laγo. 'the dog_NOM is chasing the rabbit_ACC'

Procedure

The three tasks (baseline tasks and the PS task) were administrated in random order during two sessions that lasted approximately 30 minutes each. A laptop computer was used for the data collection.

In the PS task, the participant heard a sentence and s/he had to point to one of two presented pictures. One of the pictures matched the spoken sentence, while in the other picture, the two arguments were reversed (see Figure 1). The order of experimental sentences was pseudorandomized to avoid the appearance of the same characters more than three times in a row. Also, the target picture did not appear in the same side of the computer screen (right or left) in more than three consecutive sentences.

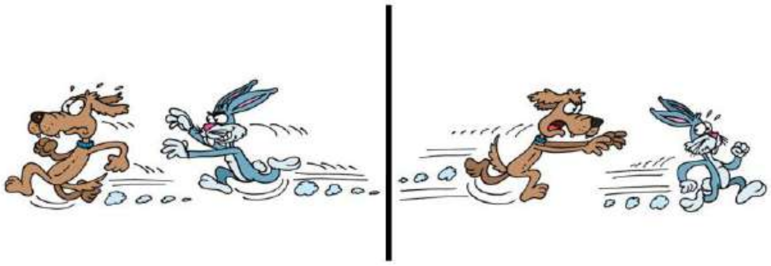


Figure 1. Sample of picture stimuli

Results

The percentage of correct responses for each group is presented in Table 2. Overall, the adult group performed better than the child group in all experimental conditions.

Table 2. Percentage of correct responses

A/α	Conditions	Child group	Adult group
1	+NP restriction, -clitic	68%	99%
2	+NP restriction, + clitic	74%	100%
3	-NP restriction, -clitic	86%	100%
4	-NP restriction, +clitic	75%	95.5%
5	plural, +NP restriction	87%	100%
6	plural, +clitic, +pro	73%	100%
7	control	97%	100%

A two-way mixed ANOVA (2 X 7) [group (children, adults) X condition (1, 2, 3, 4, 5, 6, 7)] with Greenhouse-Geisser correction revealed that there was a statistically significant interaction between groups and conditions ($F_{AxB}(4.743,438) = 12.27$, $p <.001$, $\eta^2_p = 0.15$), and a main effect of group ($F_A(4.743,438) = 15.18$, $p < .001$, $\eta^2_p = 0.17$) and condition ($F_B(1,73) = 106.75$, $p < .001$, $\eta^2_p = 0.59$), as well. Moreover, the simple effects revealed that adults performed significantly better in all experimental conditions (C1: $F(1,73) = 55.5$, $p < .001$, C2: $F(1,73) = 72.2$, $p < .001$, C3: $F(1,73) = 34.9$, $p < .001$, C4: $F(1,73) = 28$, $p < .001$, C5: $F(1,73) = 17.6$, $p < .001$, C6: $F(1,73) = 55.9$, $p < .001$, C7: $F(1,73) = 4.84$, $p = .03$).

Children had a ceiling performance in *C7* (control condition) (97%) and a high performance in C*3* (86%) and *C5* (87%). Their performance was lower in the other conditions: C*1* (68%), *C2* (74%), *C4* (75%), and *C6* (73%).

A repeated measures one-way ANOVA with Greenhouse-Geisser correction revealed significant differences between conditions ($F(4.759,176.087) = 13.67$, $p < .001$, $\eta^2_p = 0.27$). Multiple comparisons using Bonferroni correction showed that children performed significantly better in C*3* than in *C1* (M.D.: 1.08, $p = .006$, 95% C.I. [0.21, 1.95]) and C*2* (M.D.: 0.74, $p = .014$, 95% C.I. [0.09, 1.38]). Significant differences were also found between *C5* and *C6* (M.D.: 0.79, $p = .04$, 95% C.I. [0.02, 1.56]), *C1* (M.D. 1.105, p = .00, 95% C.I. [0.42, 1.79], *C2* (M.D.: 0.76, $p = .022$, 95% C.I. [0.06, 1.46]). The control condition significantly differed from all experimental conditions ($p < .05$).

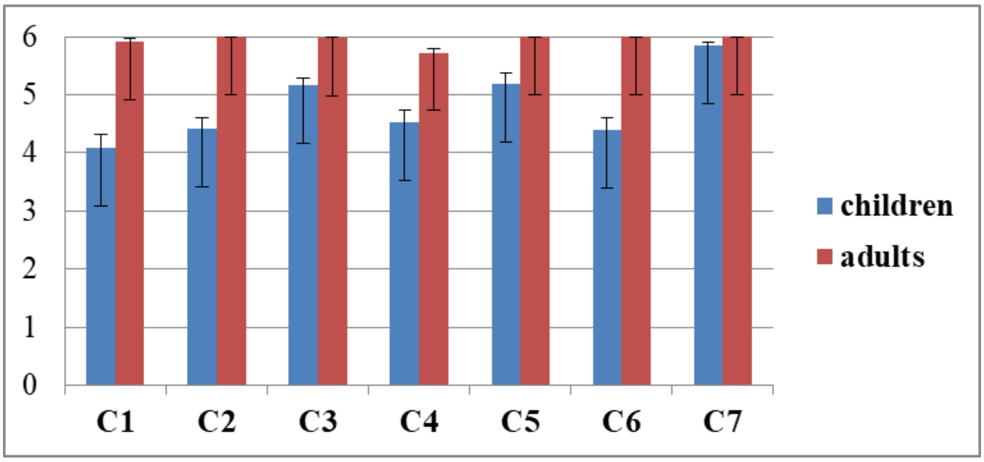


Figure 2. Error bar graphs for all experimental conditions

The adult group had ceiling performance in all experimental conditions. However, lower performance was noticed in *C4* (95.5%). Comparing the means of correct responses for each condition, a repeated measures one-way ANOVA with Greenhouse-Geisser correction revealed significant differences between conditions ($F(6,216) = 8.66$, $p < 0.001$, $\eta^2_p = 0.19$). Multiple comparisons using Bonferroni correction showed that *C4* differed significantly from *C2* (M.D.: -0.27, *p* = .17, 95% C.I. [-0.51, -0.28]), *C5* (M.D.: -0.27, *p* = .17, 95% C.I. [-0.51, -0.28]), *C6* (M.D.: -0.27, *p* = .17, 95% C.I. [-0.51, -0.28]), and the control condition (M.D.: -0.27, *p* = .17, 95% C.I. [-0.51, -0.28]).

Discussion

The results of this study confirmed Friedmann *et al*.'s (2009) claim that lexical NP restriction in both the moved and the intervening DP causes difficulties to children's comprehension of ORCs (*C3* vs. *C1*, *C2*). Furthermore, our results indicate that apart from lexical NP restriction, the internal structure of the two DPs also facilitates children's comprehension of ORCs, in accordance with Adani *et al*. (2010), Belletti *et al*. (2012), and Varlokosta *et al*. (2015). Children performed higher on the number mismatch condition compared to the number match conditions (*C5* vs. *C1*, *C2*).

Interestingly, the presence of a resumptive clitic in the original merge position of the relative clause did not facilitate children's performance regardless of the presence or absence of lexical NP restriction in the moved and intervening DP; performance on *C2* was not significantly different from performance on *C1* and the same was true for performance on *C4* compared to performance on *C3*. These results indicate that the absence of lexical NP restriction in either the moved or the intervening DP is more decisive for the reduction of intervention effects than the presence of a resumptive pronoun in the original merge position of the relative clause. This is consistent with findings in Friedmann *et al*. (2009) regarding the presence of resumptive pronouns in Hebrew. However, we should point out that although the difference between *C3* and *C4* did not reach statistical significance in the child group, the presence of a resumptive clitic in *C4* caused poorer performance in the adult group as well, which proved to be statistically significant compared to performance on all other conditions. At this point, we cannot provide an explanation for the deteriorating effect that the resumptive clitic had, particularly on the condition without lexical NP restriction. Future research should further investigate the effect that resumptive pronouns have in intervention phenomena in child language.

Our last observation concerns the effect that an arbitrary pro intervening subject had in children's comprehension of ORCs. The statistically significant difference between *C5* and *C6* confirms the conclusion reached in Adani *et al*. (2010) and Varlokosta *et al*. (2014) that lexical NP restriction in both DPs is not the only factor that affects children's performance in structures that induce intervention effects. The use of an arbitrary pro, which does not have lexical NP restriction, in the intervening subject position of the relative clause resulted in lower performance compared to the use of a DP with lexical NP restriction and number mismatch. The rather poor performance on the condition with an arbitrary pro subject is not compatible with Friedmann *et al*.'s (2009) results from Hebrew and indicates that the requirement of featural disjointness of strict RM cannot be solely satisfied by the non-lexical or null nature of the intervening subject. Friedmann *et al*. (2009) raise also the question whether the null nature of the intervening subject per se or the difference in agreement features between the moved DP and the pro subject intervener (which shows up in the verbal morphology) play a role in the resolving intervention. Our results indicate that featural mismatch on the inflection of the verb is not a sufficient morphological cue to resolve intervention effects. Clearly, further research is necessary in order to determine the role of lexical NP restriction and featural dissimilarity in intervention effects in child language.

Conclusion

This study examined the comprehension of ORCs by Greek-speaking children, aged 3;1-4;11 years, considering the effect that different kinds of intervening subjects and moved elements may have on intervention effects. Our results showed that absence of lexical NP restriction in the moved DP and number mismatch between the moved and the intervening DP improve children's comprehension of ORCs. The presence of an arbitrary pro and featural mismatch on the inflection of the verb is not a sufficient morphological cue to resolve intervention effects neither is the presence of a resumptive clitic in the original merge position of the relative clause.

References

Adani, F., van der Lely, H. K., Forgiarini, M., & Guasti, M. T. (2010). Grammatical feature dissimilarities make relative clauses easier: A comprehension study with Italian children. *Lingua*, *120*, 2148-2166.

Belletti, A., Friedmann, N., Brunato, D., & Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua*, *122*, 1053-1069.

Crain, S., McKee, C., & Emiliani, M. (1990). Visiting relatives in italy. In L. Frazier & J. d. Villiers (eds.), *Language processing and language acquisition* (pp. 335-356). Dordrecht: Kluwer.

de Villiers, J. G., Tager Flusberg, H. B., Hakuta, K., & Cohen, M. (1979). Children's comprehension of relative clauses. *Journal of Psycholinguistic Research*, *8*, 499-518.

Dunn, L., & Dunn, L. (1981). *Peabody picture vocabulary test-revised.* Circle Pines, MN: American Guidance Service.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, *119*, 67-88.

Raven, J. C. (1998). *The coloured progressive matrices*. Oxford, UK: Oxford University Press.

Rizzi, L. (1990). *Relativized minimality.* Cambridge, MA: MIT Press.

Rizzi, L. (2001). Reconstruction, weak island sensitivity, and agreement. In C. Cecchetto, G. Chierchia & M.T. Guasti (eds.), *Semantic interfaces: Reference, anaphora and aspect* (pp. 145-176). IL: CSLI.

Sideridis, G., Antoniou, F., Simos, P., & Mouzaki, A. (2015). *Raven's colour progressive matrices, Greek standardization*. Athens: Topos Publishers.

Simos, P. G., Sideridis G. D., Protopapas A., & Mouzaki, A. (2011). Psychometric characteristics of a receptive vocabulary test for Greek elementary students. *Assessment for Effective Intervention*, *37*, 34-49.

Sheldon, A. (1974). The role of parallel functions in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behaviour*, *13*, 272-281.

Varlokosta, S., Nerantzini, M., & Papadopoulou, D. (2014). Comprehension asymmetries in language acquisition: A test for Relativized Minimality. *Journal of Child Language, 42*, 618-661.

Appendix A.

Table 3. Children's age and raw scores in Raven CPM and PPVT.

Participants	Age	Raven	PPVT	Participants	Age	Raven	PPVT
C1	4;9	13	23	**C20**	3;5	10	49
C2	4;0	13	21	**C21**	4;9	12	67
C3	4;5	11	69	**C22**	4;8	7	61
C4	3;11	15	49	**C23**	4;8	12	90
C5	4;9	10	74	**C24**	4;6	15	85
C6	4;5	10	32	**C25**	4;5	5	60
C7	3;6	10	79	**C26**	4;9	16	12
C8	3;6	10	18	**C27**	4;5	11	92
C9	3;7	16	68	**C28**	4;1	7	68
C10	3;2	12	20	**C29**	4;5	12	61
C11	3;2	12	29	**C30**	4;11	17	80
C12	3;9	13	70	**C31**	4;11	11	49
C13	3;4	8	20	**C32**	4;11	13	77
C14	3;7	12	75	**C33**	4;10	4	49
C15	3;11	14	46	**C34**	4;7	15	68
C16	3;4	13	48	**C35**	4;7	8	42
C17	3;5	12	56	**C36**	4;10	11	45
C18	3;5	13	50	**C37**	3;6	11	21
C19	3;1	9	19	**C38**	3;3	7	18

Appendix B

List of sentences used in the PS task

1. o 'scilos cini'γa to la'γo 'the dog_ NOM is chasing the rabbit_ACC'
2. o kro'koðilos ci'tazi ton ipo'potamo 'the crocodile_ NOM is looking at the hippo_ACC'
3. o 'líkos ða'goni to 'sciuro 'the wolf_ NOM is biting the squirrel_ACC'
4. o pa'pus fotoɣra'fízi to 'nafti (The grandpa_NOM is photographing the sailor_ACC'
5. o 'adras çere'tai to ɟa'tro 'the man_NOM is waiving to the doctor_ACC'
6. o foto'ɣrafos tra'vai ton asti'nomo 'the photographer_ NOM is pulling the policeman_ACC'

To cite this article: Alexandri, C., Apotolopoulou, S., Drobanikou, A., Kotroni, E., & Varlokosta, S. (2017). The acquisition of object relative clauses in Greek: The role of lexical NP restriction and featural mismatch. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 7-13). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Early word learning by Turkish-Dutch bilinguals: Cognitive or linguistic dominance?

N. Feyza Altinkamiş[1], F. Hülya Özcan[2]
feyza.altinkamis@ugent.be, fozcan@anadolu.edu.tr

[1]Ghent University, [2]Anadolu University

Abstract. Maguire, Hirsh-Pasek, and Golinkoff (2006), following suggestions by Gentner and Boroditsky (2001) and Snedeker and Gleitman (2004), have offered to develop a comprehensive approach to early word learning instead of only focusing on syntactic classes, such as noun or verb. According to them, any word which represents perceptually accessible concepts, rather than only nouns or verbs, can be learned earlier than other word categories with abstract or relational concepts. Taking a broader view, Maguire *et al.* (2006) propose an amalgamation, called the SICI (Shape, Individuation, Concreteness, Imageability) continuum to understand the concepts of words acquired early by children. Xuan and Dollaghan (2012) investigated the perceptual and cognitive characteristics of the first fifty words in Mandarin-English bilingual children's early lexicon with regard to the features of the SICI continuum and concluded that cognitive and perceptual constraints were not salient enough to explain children's early word learning. In line with this background and extending the results of Xuan and Dollaghan's (2012) study involving Mandarin-English bilingual children, we investigated the nature of the most frequent words in Turkish-Dutch bilingual children's early lexicon in terms of the SICI features. The data in this study were collected using the Turkish and Dutch adaptations of the CDI from 48 Turkish-Dutch bilingual children in Flanders. The most frequent words in the children's Dutch and Turkish lexicon were determined and were grouped in line with the features of the SICI continuum. The results shows similarity with the findings of Xuan and Dollaghan (2012), indicating not only the effect of perceptual and cognitive mechanisms but also the interaction of linguistic features on early word learning.

Keywords: SICI continuum, Turkish-Dutch, early lexical development, reference, predicate, grammar

Introduction

Related literature on monolingual children's language development reveals some common developmental patterns in the composition of the early lexicon. These have also been confirmed in various studies of different monolingual populations. Of these, the most commonly found pattern among different languages is the trajectory from reference (nouns) to predication (verbs and adjectives), to grammar (closed-class items). Bates, Marchman, Thal, Fenson, Dale, Reznick, and Reilly (1994) studying English speaking children, Stolt, Haataja, Lapinleimu, and Lehtonen (2008) studying Finnish speaking children, Caselli, Bates, Casadio, Fenson, Sanderl, and Weir (1995) studying Italian children, Kern (2007) studying French speaking children, and Eriksson and Berglund (1999) studying Swedish speaking children, provide parallel data through different adaptations of the MacArthur Communicative Development Inventory (CDI). Bassano (2000) and Türkay (2009) presented similar results using naturalistic data in French and Turkish, respectively. In Bassano's longitudinal study, a shift from nouns to predicates and to grammar was also observed in Türkay's naturalistic longitudinal study that involves five Turkish children. Although all the above-mentioned languages show different language characteristics, the developmental trajectory is common in the early lexical composition of children in all these languages. The developmental change from common nouns to predicates, and then to closed-class items, is also in line with Gentner's (1982) and Gentner's and Boroditsky's (2001) revised cognition-based approach). According to the universal noun advantage view proposed by Gentner (1982), children's early lexicon is made up of nouns, representing concrete objects. Verbs and verb-like items are acquired later as they require a cognitively complex task for children to accomplish.

Maguire *et al*. (2006), following suggestions by Gentner and Boroditsky (2001) and Snedeker and Gleitman (2004) have offered to develop a comprehensive approach into early word learning, instead of only focusing on syntactic classes such as noun or verb. According to them, neither nouns nor verbs, but any word which represents perceptually accessible concepts can be learned earlier than other word categories with abstract or relational concepts. Taking a broader view, Maguire *et al*. (2006) propose an amalgamation to understand the concepts of the early acquired words by children, called the SICI continuum. The SICI is an acronym representing four factors; shape, individuation, concreteness and imageability. According to Maguire *et al*. (2006), it is possible to develop a broad perspective into children's early word learning using SICI, where not only one feature of the continuum is considered, because related literature is full of different terms, and operational definitions are rather opaque. They also add that there is indefiniteness among the factors in terms of the weight they carry in the continuum. All four factors seem equally important in order to grasp the meaning of how the SICI continuum of concepts works (Figure 1):

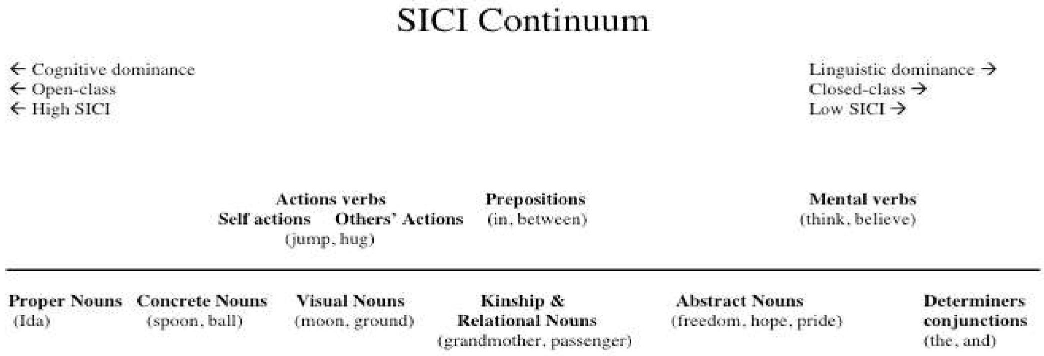


Figure 1. The SICI continuum (Maqurie *et al.*, 2006:57)

- S is for shape: some verbs are linked to certain shapes and result in fast mapping in children's early word learning in experimental studies. Objects are definitely representative of the shape factor, but verbs can also be put into this scale in line with how much shape consistency they show. For example, *dancing* is easier in terms of shape than *thinking* (Maguire *et al*., 2006:18).

- I is for individuation: Maguire *et al*. (2006) here refer to Gentner and Boroditsky's Division of Dominance (2001). In their continuum, there are two ends. At the cognitive end, there are objects which are easy to individuate, namely open-class-category words, and at the linguistic continuum, there are items which are mostly related to language, namely the closed-class category. Gentner and Boroditsky (2001) propose that verbs are at the more difficult end of the continuum and they are harder to acquire than nouns.

- C is for concreteness and I is for imageability: Maguire *et al*. (2006) say that concreteness and imageability have been used interchangeably in the literature. For 'imageability' and 'concreteness', they refer to the definition given by Paivio, Yuille, and Madigan (1968). According to them, 'imageability' is to what extent the word is easy to link to a sensory mental image and 'concreteness' is the ability to see, hear and touch something.

The authors sum up the interaction of the four factors as follows:

> 'words can be thought of as falling on a continuum that characterizes the reliability and consistency of their shape. The ease with which they can be distinguished from other items in the scene (individuability); whether they can be observed in the world at all and whether they are manipulable (concreteness); and how readily they yield a mental image for adults (imageability)' (Maguire *et al*., 2006:22-23).

Against this background, Xuan and Dollaghan (2012) investigated the perceptual and cognitive characteristics of the top 50 words in Mandarin-English bilingual children's early lexicon with regard to the features of the SICI continuum and concluded that cognitive and perceptual constraints were not salient enough to explain children's early word learning. In our study, we would like to concentrate on a new bilingual group, Turkish-Dutch bilinguals, with a similar perspective to that of Xuan and Dollaghan (2012) and formulate the following research questions:

1. To what extent are the most frequent words in the Turkish-Dutch bilingual children's early lexicon compatible with the features of the SICI continuum?
2. Are there commonalities between the Turkish and Dutch lexicon of these bilingual children in terms of the SICI features?

Method

Participants

The parents of 48 children participated in the study. They were given the Turkish and Dutch version of the MacArthur-Bates Communicative Development Inventory: Words and Sentences (M-CDI-II) and once they were given some explanations about completing the inventories, they were invited to fill these in independently. Seventeen children were from one-parent-one-language families, and 31 children were from families with monolingual Turkish parents. The children's ages ranged from 1;05 to 2;11 (M=27.2 months). Twenty-four of the children were girls and 24 of them were boys.

Data Collection Tools

M-CDI-II was used as a data collection tool in the study. CDI has two forms: CDI-I Infant Form (Word and Gestures) addressing children from 8 to 16 months, and CDI-II Toddler Form (Words and Sentences) addressing children from 16 to 30/36 months. In this study, we only focus on CDI-II Toddler Form. Table 1 gives detailed information about the Turkish (CDI-TR) (Aksu-Koç, Küntay, Acarlar, Maviş, Sofu, Topbaş, & Turan, 2009) and the Dutch (CDI-DT) adaptation of CDI (Zink & Lejaegere, 2002).

Table 1. CDI Toddler Form Description

	Toddler Form	
	CDI-TR	CDI-DT
Lexical categories	21	22
Total words	711	702

Data Analysis

To perform the frequency analysis, we tallied the percentage of children who used the words in Dutch and Turkish and prepared a list of the most frequent Turkish and Dutch words. Due to the ties at the 50th rank, 52 Turkish words and 56 Dutch words were categorized in the children's lexicon.

Table 2. Most frequent 52 Turkish words and their CDI categories

word	**English**	**CDI category**	**%**
baba	'daddy'	people	97
anne	'mummy'	people	93
dede	'grandpa'	people	89
balon	'balloon'	toys	77
su	'water'	food and drink	77
abi	'brother	people	77
alo	used answering the phone	games and routines	77
kaka	used for the need to go to toilet	games and routines	77
araba	'car'	vehicles	75
bebek	'baby'	people	75

hop	sound effect of jumping	sound effects/animals	72
ham	sound effect of eating	sound effects/animals	70
bebek	‘babydoll’	toys	70
abla	‘elder sister’	people	68
baybay	‘good-bye’	sound effects/animals	68
muz	‘banana’	food and drink	66
uf	sound effect of being in pain	sound effects/animals	62
at	‘horse’	animals	62
süt	‘milk’	food and drink	62
ayakkabı	‘shoes’	clothing	62
kendi ismi	the child's own name	people	62
teyze	‘maternal aunt’	people	62
düt	sound effect of a car	sound effects/animals	60
köpek	‘dog’	animals	60
kuş	‘bird’	animals	60
ayak	‘foot’	body parts	60
*otur**	‘sit down’	action words	60
şişt	sound effect of warning	sound effects/animals	58
top	‘ball’	toys	58
çay	‘tea’	food and drink	58
burun	‘nose’	body parts	58
mama	‘food’	food and drink	58
yok	‘no/absent’	games and routines	58
balık	‘fish’	animals	56
kedi	‘cat’	animals	56
ekmek	‘brood’	food and drink	56
et	‘meat’	food and drink	56
çorap	‘socks’	clothing	56
el	‘hand’	body parts	56
göz	‘eye’	body parts	56
çiş	used for the need to go to toilet	games and routines	56
*ver**	‘give’	action words	56
acı	‘bitter’	descriptive words	56
cıss	warning of danger	sound effects/animals	54
elma	‘apple’	food and drink	54
ev	‘house’	places to go	54
anneanne	‘maternal grandma’	people	54
babaanne	‘paternal grandma’	people	54
evet	‘yes’	games and routines	54
acı (canı) *	‘to have pain’	action words	54
at *	‘throw away’	action words	54
op *	‘kiss’	action words	54

Results

We first highlighted the nouns in bold and then the verbs with an asterisk. In the list, 23 nouns were written in bold and 5 verbs were marked with an asterisk in the Turkish lexicon of the children (see Table 2) and 34 nouns were made bold and five verbs were marked with an asterisk in the Dutch lexicon of the children (see Table 3).

Table 3. Most frequent 56 Dutch words and their CDI categories

word	English	CDI category	%
waf waf	‘woof woof’	sound effects/animals	87
hallo	used answering the phone	sound effects/animals	83
miauw	‘miaow’	sound effects/animals	77
mjam	sound effect of ‘delicious’	sound effects/animals	77

auto	‘car’	vehicles	77
bal	‘ball’	toys	72
appel	‘apple’	food and drink	72
au ai	sound effect of being in pain	sound effects/animals	68
baby	‘baby’	people	66
dada doen	‘good-bye’	games and routines	66
ballon	‘balloon’	toys	64
koekje	‘cookie’	food and drink	64
mama	‘food’	food and drink	64
beu boe	‘moo’	sound effects/animals	62
fiets	‘bicycle’	vehicles	62
banaan	‘banana’	food and drink	60
vroem broem	‘vroom’	sound effects/animals	58
ja	‘yes’	games and routines	58
boek	‘book’	toys	56
water	‘water’	food and drink	56
oog	‘eye’	body parts	56
hallo	‘hello’	sound effects/animals	56
kwak kwak	‘quack’	sound effects/animals	54
oh ooh	sound effect of surprise	sound effects/animals	54
jas	‘jacket’	clothing	54
schoen	‘shoe’	clothing	54
oor	‘ear’	body parts	54
papa	‘daddy’	people	54
kip	‘chicken’	animals	52
koe	‘cow’	animals	52
blokken	‘blocks’	toys	52
eten	‘food’	food and drink	52
frieten	‘fries/chips’	food and drink	52
soep	‘soup’	food and drink	52
luier/pamper	‘diaper/nappy’	clothing	52
neus	‘nose’	body parts	52
water	‘water’	outside things	52
oma	‘grandma’	people	52
nee	‘no’	games and routines	52
beh	‘baa’	sound effects/animals	50
hondje	‘dog’	animals	50
aardappel/patatjes	‘potatoes’	food and drink	47
aardbeien	‘strawberry’	food and drink	47
brood	‘bread’	food and drink	47
chips	‘chips/crisps’	food and drink	47
melk	‘milk’	food and drink	47
naam van kind	the child's name	people	47
kiekeboe	‘peekaboo’	games and routines	47
kat/poesje	‘cat’	animals	45
haar	‘hair’	body parts	45
hand	‘hand’	body parts	45
telefoon	‘telephone’	small household items	45
bad	‘bath’	furniture and rooms	45
bed	‘bed’	furniture and rooms	45
opa	‘grandpa’	people	45
*eten **	‘to eat’	action words	45

When we group the most frequent words into the CDI categories, we see that lexical categories in common nouns were: animals, toys, food and drink, clothing, body parts, small household objects, furniture, rooms and vehicles. Action names and adjectives were included in predicates. Next, pronouns, question words, prepositions (in Dutch), articles and quantifiers, auxiliary and modal verbs (in Dutch) and connecting words were grouped in the closed-class category. Lastly, names for people, games, routines and sound effects made up social words (Bates *et al*., 1994; Caselli *et al*., 1999; Kern, 2007; Stolt *et al*., 2008). As illustrated in Tables 2 and 3, there were no words from the CDI closed-class category, but children used words from people, games, routines and sound effects categories. Next, we followed Xuan and Dollagan (2012) for the SICI features analysis. Nearly half of the words (44.23 %) in the children's Turkish lexicon constituted the CDI noun category and were located towards the high end of the SICI continuum in terms of shape, individuation, concreteness and imageability, such as *muz* 'banana' and *kuş* 'bird'. Categories *Games & Routines* and *Sound Effects & Animal Sounds* were also found to be important (23%) in the Turkish lexicon. In the children's Dutch lexicon, more than half of the words (57%) belonged to the CDI noun category. The categories *Games & Routines* and *Sound Effects & Animal Sounds* constituted 26% of the words. Therefore, we can say that although words from the high end of the SICI features are seen, they do not refer to the SICI features available in the children's most frequent 50 words.

As for verbs, in the children's Turkish lexicon, there are only five verbs (*otur* 'sit down', *ver* 'give', *acı* 'to have pain', *at* 'throw away' and *öp* 'kiss'. In their Dutch lexicon, there is only one verb *eten* 'to eat'. These verbs in the children's early lexicon can be regarded as semantically heavy according to Xuan and Dollaghan (2012), carrying specific meanings, but not light verbs.

Discussion

In this study, we investigated the nature of most frequent words in Turkish-Dutch bilingual children's early lexicon in terms of the SICI features. Concentrating on the children's early lexicons individually, we can say that our findings support the results by Xuan and Dollaghan (2012). Most frequent words in the children's Dutch and Turkish lexicons are made up of words representing common nouns, but there are also a considerable number of words which do not match the SICI features, such as *Games & Routines* and *Sound Effects & Animal Sounds*. For example, when talking about the most frequent words list in terms of the SICI continuum, no emphasis is given to the people category in the study by Xuan and Dollaghan (2012), but kinship terms representing the *People* category are emphasized in the SICI continuum (see Figure 1). There is a discrepancy among the studies about the content of the *Common Nouns* category. The *People* words are sometimes grouped under the *Common Nouns* category in the CDI, while they are excluded in other studies. However, we believe that the words in the *Common Nouns* category highlight the importance of children's immediate environment and culture-specific characteristics. In the children's Turkish lexicon, there are more words from the *People* category than their Dutch lexicon. This is also the pattern observed in the study by Rinker, Budde Spengler, and Sachse (2016) with Turkish-German bilingual children. Two reasons are given for this. One is that there are more words in the Turkish language than in German in terms of family connections, and the other is that this reflects the traditional Turkish family model. In addition to that, we believe that the words in the *People* category fit the SICI features well. For example, *grandpa-grandma* refer(s) to shape, easy individuation, concreteness and imageability. Another most frequent CDI category is food and drinks, which shows the importance of basic and prominent needs in this age group.

Conclusion

We can conclude that nearly half of the most frequently used words in the Turkish-Dutch bilingual children's early lexicon are compatible with the features of the SICI continuum (44.23% in the Turkish lexicon and 57% in the Dutch lexicon) but there are a considerable number of words that are not represented by the SICI continuum. In general, we can conclude that the children's early lexicon is based on words referring to cognitive and perceptual constraints, as proposed by Gentner (1982) in terms of the SICI features, but only focusing on these perspectives may lead us to ignore the other words in the children's early lexicons. Other factors that shape the structure of children's early lexicon are their social environment and their interactions.

References

Aksu-Koç, A., Küntay, A., Acarlar, F., Maviş, İ., Sofu, H., Topbaş, S., & Turan, F. (2009). The Assessment and evaluation of early lexical and grammatical development in Turkish: The Turkish Communicative Development Inventories, TIGE-I and TIGE-II. *Final report of Project no: 107K058, submitted to Turkish Scientific and Technological Research Foundation.*

Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, *27*(3), 521-59.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P. Reznick, S., Reilly, J.S., & Hartung, P.J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21*, 85-123.

Caselli, M., Bates, E., Casadio, P., Fenson, L., Sanderl, L., & Weir J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, *10*(2), 159-200.

Eriksson, M., & Berglund, E. (1999). Swedish early communicative development. *First Language*, *19*(1), 55-90.

Gentner, D. (1981). Some interesting differences between nouns and verbs. *Cognition and Brain Theory*, *4*, 161-178.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (ed.), *Language development: Vol. 2. Language, thought, and culture* (pp.301-333). Hillsdale, NJ: Erlbaum.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman & S. Levinson (eds.), *Language acquisition and conceptual development* (pp. 215-256). Cambridge, UK: Cambridge University Press.

Kern, S. (2007). Lexicon development in French-speaking infants, *First Language*, *27*(3), 227-250.

Maguire, M. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2006). A unified theory of word learning: Putting verb acquisition in context. In K. Hirsh-Pasek & R. M. Golinkoff (eds.), *Action meets words: How children learn verbs* (pp. 364-391). New York, NY: Oxford University Press.

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*, 1-25.

Snedeker, J., & Gleitman, L. (2004). Why is it hard to label our concepts. In G. H. Waxman (ed.), *Weaving a lexicon*. Cambridge, MA: MIT Press.

Stolt, S., Haataja, L., Lapinleimu, H., & Lehtonen, L. (2008). Early lexical development of Finnish children - a longitudinal study. *First Language*, *28*(3), 259-27.

Rinker, T., Budde Spengler, N., & Sachse, S. (2016). The relationship between first language (L1) and second language (L2) lexical development in young Turkish-German children. *International Journal of Bilingual Education and Bilingualism*, *20*(2), 1-16.

Türkay, F. (2009). The developmental paths in Turkish children's early lexicon. *Çukurova University, Journal of Social Sciences*, *18*(2), 415-426.

Xuan, L., & Dollaghan, C. (2012). Language-specific noun bias: evidence from bilingual children. *Journal of Child Language*, *40*(5), 1057-1075.

Zink, I., & M. Lejaegere. 2002. *N-CDIs: Lijsten voor communicatieve ontwikkeling. Aanpassing en hernormering van de MacArthur CDIs van Fenson et al*. Acco, Leuven (Belgium)/Leusden (Netherlands).

To cite this article: Altinkamiş, N. F., & Özcan, F. H. (2017). Early word learning by Turkish-Dutch bilinguals: Cognitive or linguistic development? In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 14-20). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Phonologically induced syntactic transfer in L3 acquisition of English: Evidence from late sequential Russian-German bilinguals

Tanja Angelovska
tanja.angelovska@sbg.ac.at

University of Salzburg

Abstract. This study aims to investigate whether late sequential Russian-German bilinguals acquiring English as L3 will display negative syntactic transfer of the verb-second (V2) in speaking and writing and to what extent, as well as whether this transfer behavior is mediated by other factors, such as proficiency, dominance, and phonological similarity. The participants were 13 adult L3 learners (mean age = 29.6), whereby L2 German and L3 English were acquired subsequently. Two groups of different proficiency levels (elementary and intermediate) with different dominant languages used daily were formed based on *Oxford Quick Placement Test* and *German Use* scores. Data was elicited by using a time-controlled written narration task (with given adverbs as prompts) and a spoken story telling task. *Obligatory Occasion Analysis (OOA*) was used to account for the trajectory of the negative transfer appliance and a classification of the verb-preceding element according to phonological similarity between the verb-preceding initial adverbial in L3 and the corresponding phonologically similar L2 'doublet'. The results revealed very small percentage of negatively transferred V2. The transfer from the L2 appears to be triggered phonologically at a representational level but realized at a performance level in the L3. The negative syntactic transfer from the L2 was conditioned by language dominance, proficiency, and phonological similarity. Beginners transfer more in speaking than in writing, and beginners transfer more in speaking than intermediate learners. Similarly, beginners with high scores of *German Use* produce more V2 sentences in both speaking and writing tasks than those with low scores. The results offer new insights – relevant for L3 classroom instruction.

Keywords: third language acquisition, phonology, syntactic transfer, proficiency, language dominance, verb-second, phonological similarity

Introduction

The existing transfer models in third language acquisition (L3A) are based on explaining which of the prior languages will take over the leading role. Some studies consider that only one language (either the L1 or the L2) will have the privileged role. Others consider the possibility that both languages can be the source for transfer in L3A (e.g., the Cumulative Enhancement Model (CEM) and the Typological Proximity Model (TPM)) (cf. Angelovska & Hahn, 2017 for a most recent overview).

These models mainly account for the mental representation of the prior languages in L3A by using comprehension or production data (either spoken or written data) with a focus on either elementary or intermediate learners. What remains still unanswered is whether L3 learners will display negative syntactic transfer in speaking and writing and whether this is mediated by other factors, such as proficiency, dominance, and phonological similarity (cf. Angelovska, 2017b for the role of prior languages).

Theoretical background

The role of phonology in existing transfer models in L3A

Although the literature on phonological acquisition in bilingual children is abundant (cf. Babatsouli & Ingram, 2015), the number of studies focusing on the role of phonology (in particularly for transfer) in the L3A field is still limited. TPM is the only model that takes into consideration the role of phonology in L3 transfer. It predicts that one of the two previous languages will be transferred completely at the initial stages (cf. Rothman, 2013). TPM gives plausible arguments that L3 transfer can be both facilitative and non-facilitative, and proposes that it is selected by factors related to comparative structural similarity determined by the organization of the internal parser. According to the TPM, the linguistic parser operates in an implicational hierarchy of the following four linguistic cues: the lexicon, phonological cues, functional morphology, and syntactic structure.

Concrete results regarding phonological transfer come from a study on the production of voiceless stops in stressed onset position in L3 Spanish showing that the L2 status is the determining factor in the selection of a source language (Llama, Cardoso, & Collins, 2010).

Transfer, language proficiency and dominance

Studies reporting on the role of language proficiency in L3 transfer show mixed results. According to Falk and Bardel (2010a), the L2 language will no longer be the default for transfer if the learner in the L2 reaches high proficiency. However, in a later study they reported that there is an L2 transfer into L3A even at an intermediate level (Falk & Bardel, 2010b). Similarly, there are no studies investigating the role of language dominance for predicting L3 transfer. An exception is a study by Lloyd-Smith Gyllstad and Kupisch (2016), which analyzed global accent[1] and showed that heritage speakers seem to benefit from bilingual experience and structure-based transfer can be overpowered by high proficiency.

Although there are no studies on verb second transfer in L3A that considered the factor ‘language dominance’ as a predictor for source of phonologically induced syntactic transfer, results from Kupisch (2007) offer some valuable insights that can be used as a starting point. She found that the two languages of a bilingual influence each other, and that both dominance and language internal factors are decisive in determining influence.

Target: the verb-second phenomenon

V2 refers to the phenomenon where the finite verb is required to appear in the second position of a declarative main clause preceded by a single arbitrary constituent (cf. Haider, 2010). The initial sentence arbitrary constituent is not (necessarily) the subject. In German – a V2 language – the subject and the verb will be inverted, i.e. the verb appears before the subject in tropicalized main clauses. Unlike German (1) but like English (2), Russian (3) does not allow either VSO or XVSO orders, though XVOS order is possible. In Russian if the subject is preceded by a verb, the subject must appear in sentence final position.

(1) German with English translation:

Zuerst dachte ich er hat Interesse.

* ‘At first thought I he is interested.’

Adv V S O

(2) English:

Yesterday I kissed him.

Adv S V O

(3) Russian with English translations:

В прошлом году мой друг построил дом возле озера.

‘Last year my friend built a house near a lake.’

Adv S V O

Дом возле озера мой друг построил в прошлом году.

[1] See Angelovska (2012) for global accent in L2 pronunciation.

* ‘A house near a lake my friend built last year.’

O S V Adv

В прошлом году построил дом возле озера мой друг.

* ‘Last year built a house near a lake my friend.’

Adv V O S

The V2 phenomenon has not been the exclusive focus of existing studies on transfer in L3A, with some exceptions such as Sağin Şimşek (2006) and Bardel and Falk (2007). Sağın Şimşek (2006) presented 38 negative transfer cases of a total number of 465 sentences in which Turkish-German bilinguals adopted the German V2 mechanism for both main and subordinate clauses, placing the finite verb in a second position in L3 English. She hypothesized that the typological similarities between German and English were the reason for the syntactic transfer. Bardel and Falk (2007) found differences between the English as L2 (EN group) and Dutch as L2 group in the transfer of negation stating clearly that ‘none of the learners in the EN group systematically transfers the placement of negation of his or her L1, although the L1 shares the V2 rule with the L3’ (Bardel & Falk, 2007:479).

This study attempts to explain the differential negative (L2 to L3) transfer outcomes by considering the dominant language used daily (either L2 German or L1 Russian) and by accounting for the phonologically triggered L2 activation in L3 (English) transfer.

Hypotheses

The following hypotheses guide this study:
(1) H1: L3 beginners of English will display more occurrences of phonologically induced negative syntactic transfer of V2 from L2 German in L3A of English than L3 intermediate learners of English will do at both written and spoken tasks.
(2) H2: L3 beginners of English will display more occurrences of phonologically induced negative syntactic transfer of V2 from L2 German in L3A of English at the spoken than at the written task.
(3) H3: If there is a negative interlanguage transfer of V2 from L2 German in L3A of English it is dependent on language dominance and property-by-property language structural similarities triggered phonologically.

Method

Participants

The participants were 13 adult learners of L3 English with L1 Russian and L2 German at different elementary and intermediate proficiency levels, aged between 24 and 37 (mean= 29.6). The L3 subjects had not reported any stay in an English-speaking country. They had not reported any knowledge or use of other foreign languages, nor had they reported attending foreign language courses in languages other than English during the data collection time. They all resided in Germany.

Materials

Participants were tested with the *Oxford Quick Placement Test (OQPT)* that determined their proficiency level in L3 English. According to the OQPT results, two main groups were labeled as beginners and intermediate, whereby the first included levels A1 and A2 (*n*= 6) and the second group – the intermediate – included B1 and B2 learners (*n*=7).

In a similar fashion as Lloyd-Smith *et al*. (2016), a language use dominance scoring procedure was used for determining the dominant language (see Table 1). It provided information about the subjects’ use of the languages under consideration (German, Russian and English). We calculated a score representing the quantity and quality of their contact with German in the past and at the time of testing (henceforth, referred to as German Usage Score – GUS). Some factors were weighed more heavily

CATEGORY	TYPE OF QUESTION
Language within family	Language of mother (1 point)
	Language with mother (1 point)
	Language of father (1 point)
	Language with father (1 point)
	Language between parents (1 point)
	Language among siblings (1 point)
	Language at home (1 point)
Language quality	German schooling in Russia (2 points)
	Types of contact with German (e.g. reading, writing, media etc.) (3 points)
Current language use	Relative use of Russian vs. German (3 points)
	Language/s at work/University (1 point)
	Language in spare time (1 point)
	Number of people German is spoken with (1 point)
	Relationship to people German is spoken with (work, private) (1 point)
Language experience in Germany	Number of years living in Germany (2 points)
	Number of visits per year in Germany prior to moving (2 points)

Table 1. Information used for German Use Score (max points 23 points)

than others, as they were believed to have greater influence (for example, duration of schooling in German, types of contact and duration of stays in Germany).

The data for the *written narrative task* was elicited by using a list of 5 adverbials (then, so, last year, at that night, here) as a stimulus which learners were supposed to use and narrate a memorable event from their past. The task was time–controlled (ca. 13-15 minutes).

It was not familiar to the learners and it was administered to them in class. The participants were not allowed to ask questions related to the vocabulary or the development of the story, or to use helpful tools. The activity was carried out under supervision. Learners were not given any time for preparation, but they had the list of adverbials in front of them while writing.

The data for the spoken task was elicited by using a story telling task called ‘Frog, where are you?’ based on a 24-picture cartoon. Each picture represents an event. The whole set of pictures is rich in opportunities for encoding temporal distinctions which predicted an expected use of adverbials.

Each of the 13 participants was recorded individually at a time together with an English-speaking person, during approximately 25 minutes. Before the story narration started, the participants were instructed and introduced to the task. A short small-talk sequence was included in English, without using any German or Russian to prompt their language activation in L3 English.

Prior to starting the data elicitation, both tasks had been piloted with 3 students at beginner levels in German, and further used for the study on L3 English. The tasks and given time restrictions worked out satisfactorily.

Analysis and scoring

The recordings were transcribed. Two researchers counted all instances of word orders (V2 and non-V2). The structures were analyzed in terms of correct or incorrect subject-verb placement. Imperative sentences or incomplete incomprehensible sentences were discarded from the analysis.

Obligatory Occasion Analysis (OOA) was used to examine ‘how accurately learners use specific linguistic (usually grammatical) features’ (Ellis & Barkhuizen, 2005:73) and to account for the developmental trajectory. First, samples of learner’s productions were collected. Then, obligatory occasions for the use of specific features were identified, i.e. number of all tropicalized sentences. Third, the percentage of accurate use of the feature (i.e. AdvSVO) was then calculated by means of the following formula:

$$\frac{n \text{ correct suppliance in contexts (in this case XSVO)}}{\text{total obligatory contexts (in this case total tropicalized)}} \times 100 = \text{\% accuracy (syntactic WO)}$$

In line with Ellis and Barkhuizen (2005), the criterion level of accuracy of production was set at 80% meaning that those who are below 80% are the individuals that display transfer of the V2.

Statistical comparisons to test differences between two independent groups were made by using *t-tests* for uncorrelated means, following validation for normal distribution using the Shapiro Wilk when no two values have the same value. The Pearson correlation coefficient was used to test independence between variables. For all statistical analyses, the confidence interval of difference was 95%. In what follows only statistical significances with the *p*-value lower than 0.05 will be reported.

A coding procedure for all V2 sentences by two researchers was conducted accompanied by a phonological analysis according to the *International Phonetic Alphabet* that should confirm or rule out phonologically triggered negative syntactic transfer.

Results

An overview of the whole data set and the various distributions of SVO, XSVO and XVSO reveals that only 4% of the complete data set refer to negative transfer (i.e. V2 sentences in L3). The remaining XSVO sentences are 12% and SVO are 84%. As Figure 1 displays, the difference between correct and incorrect tropicalized sentences at the beginner level in the written data set is significant with *p*-value=0.0005, at the intermediate level in the spoken data set the *p*-value is equal to 0.0195 and at the intermediate level in the written data *p*-value=0.0135.

The OOA formula was used to account for the accuracy percentages. The comparisons between written and spoken OOA and proficiency levels reveal the following results: significant differences (p<0.01) between written and spoken OOA for beginners, between OOA written beginners and OOA spoken intermediate and between OOA spoken beginners and OOA spoken intermediate. A disproportional trendline between proficiency and number of V2 occurrences could be traced, i.e. increasing proficiency and decreasing L2 transfer. The analysis of the GUS high versus OOA accuracy mode pairs showed the following significant differences (p<0.01): OOA written beginners and GUS high, OOA written intermediate and GUS high, OOA spoken beginners and GUS high, as well as OOA spoken beginners and GUS high. See Table 2 for the raw scores. Sentences with negative syntactic transfer were classified according to three categories: phonologically similar beginnings of the verb-preceding element, phonologically similar beginnings and endings of, and equivalents and/or homophones

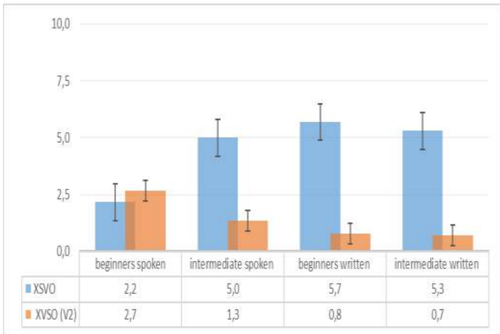


Figure 1. Mean ratings comparison between XSVO and XVSO (V2)

Phonologically similar beginnings (n= 11)	**Phonologically similar beginnings and endings (n=11)**	**Equivalents and/or homophones(n=11)**
Often impose she order in a family E: ˈɒf(ə)n G: oft / ɔft	At **night** leave the dog weg E: naɪt G: **N**ach**t** /naxt	In the small **city** stay a lot women… E: ˈsɪti G: sɪti:
In the **morning** wakes up Frosch E: ˈmɔːnɪŋ G: **mor**gen /mɔʁgən	**Then** comes owl..and **dann** bring he E: **ðɛn** G: **dan**	In dem **Moment** try the frog… E: ˈməʊmənt G: moːmɛnt

Table 3. Classification of V2 sentences according to the similarity between L2 adverb and L3 verb-preceding adverb

Learner ID	Age	L3 level acc.to OQPT	Self-ratings of language dominance	German Usage Score	OOA written accuracy in %	OOA spoken accuracy in %
1R_Alex	25	A1	Russian	12	82	50
2R_Mar	30	A1	German	13	78	17
3R_Sv	36	A2	Russian	13	83	100
4R_Olga	28	A2	Russian	16	100	50
5R_Lil	29	A2	German	15	100	40
6R_Anas	24	A2	Russian	13	100	56
7R_Svetle	36	B1	Russian	13	100	100
8R_Olgi	29	B1	Russian	16	100	67
9R_Vasili	32	B1	German	12	100	83
10R_Svet	37	B2	Russian	13	100	100
11R_Emi	25	B2	German	16	100	100
12R_Anna	30	B2	German	15	100	100
13R_Alexei	24	B2	German	16	64	64

Table 2. Participants with information about age, L3 level OQPT, self-ratings, GUS scores and OOA %

between the verb-preceding adverb in L3 English, and the corresponding L2 adverb. The classification includes all V2 sentences from both spoken and written data. In each of the three main categories, the distribution of the V2 sentences was equal (11 per category). Table 3 only illustrates 2 representatives of each category.

Discussion

Our data showed partial compatibility with the TPM: there was a transfer from the L2, but it was phonologically triggered, i.e. the preceding tropicalized element pushed the parser to predict an upcoming verb, parallel to a German sentence structure. Similarly, our data set entails cases where the ADV in the L3 had the same lexical meaning as the ADV from the L2, but they differed phonologically. The analysis of the 33 V2 sentences in the whole data set reveals that the tropicalized element in all sentences was phonologically like the L2 corresponding adverb in all cases. We cannot ignore the fact that some of these adverb pairs also display lexical equivalence, however the numerous cases where the ADV in the L3 had the same lexical meaning as the ADV from the L2, but they differed phonologically support the phonologically induced transfer from the L2, triggered at a representational level, but realized at a performance level in the L3. Although 12% (116 sentences of total 938) were AdvSVO sentences, it would be problematic to claim that this is due to positive L1 transfer from Russian, as it cannot be distinguished whether such a canonical word order ((X)SVO) is not a characteristic of a developmental stage through which all learners of a target language irrespective of their L1 go through. Similarly, if there was full transfer from the L1, we should not have faced a single occurrence of non-facilitative AdvVSO word order (from L2) in all obligatory contexts.

Comparing transfer behavior between the two modes of speaking and writing in L3A showed that beginners transfer more in speaking than in writing, and beginners transfer more than intermediate ones in speaking. An aspect of the data that deserves attention is the individual variation evident with one A2 level and two intermediate students, whereby a significant difference between the self-ratings and the GUS scores for both proficiency levels was found (p-value <0.0001), which justifies the problem of relying on self-ratings when it comes to assessment of language dominance.

We further hypothesized that language dominance will determine the source of transfer, i.e. if learners are Russian-dominant, there will not be any occurrences of L2 German negative interlanguage transfer in L3A and if learners are German-dominant, L3 learners will apply German subject-verb inversion systematically in all obligatory contexts. It was found that there are significant differences between beginners with high GUS scores and their accuracy in writing and speaking, implying that they produce V2 sentences in both speaking and writing. But, intermediate subjects who have high GUS scores do not seem to display high OOA scores in writing. These data confirmed the first part of H3 – when there was a negative interlanguage transfer of V2 from L2 German in L3A of English, it was dependent on language dominance. The second part of the H3 hypothesis—that the transfer is property-by-property and that language structural similarities are triggered phonologically—has been supported through the analysis of the V2 cases and the categorization of the verb-preceding adverbs with their phonological analyses, based on phonological similarity to the corresponding V2 adverbial, revealing that if V2 transfer in L3 was traced, it was induced phonologically (cf. Angelovska, 2017a; Angelovska & Hahn, 2012).

Conclusion

If there is negative syntactic transfer from the L2, it is phonologically triggered, i.e. the preceding tropicalized element pushes the parser to predict an upcoming verb – parallel to a German sentence structure. While we did not find uniform complete transfer from the typologically closest language, we can say that TPM has been partly confirmed because the transfer was conditioned by language dominance, proficiency, and phonological similarity.

Although our data comes from a small number of participants, the results are still important as they are based on data from both written and spoken L3 English and account for additional, so-far less explored factors. Future studies should focus not only on testing the existing L3 transfer models, but also not neglect further important factors – e.g. focus on explaining the functioning of the parser and the allocation of the cognitive resources needed to inhibit phonologically triggered transfer.

The question of how we can apply these results that tap onto the interface between phonology and syntax into L3 language instruction still needs to be addressed (see Angelovska & Hahn, 2014; Hahn & Angelovska, 2017 for first attempts).

References

Angelovska, T. (2012). *Second language pronunciation: Attainment and assessment*. Frankfurt: Peter Lang.

Angelovska, T. (2017a). (When) do L3 English learners transfer from L2 German? Evidence from spoken and written data by L1 Russian speakers. In T. Angelovska & A. Hahn (eds.), *L3 syntactic transfer: Models, new developments and implications*. [BAP Vol. 5] (pp. 195-222). Amsterdam: John Benjamins.

Angelovska, T. (2017b). Beyond instructed L2 grammar acquisition: Theoretical insights and pedagogical considerations about the role of prior language knowledge. Special Issue 'Instructed L2 Grammar Acquisition and Beyond'in *Studies in Second Language Learning and Teaching*, *7*(3), 397-419.

Angelovska, T., & Hahn, A. (eds.) (2017). L3 syntactic transfer - an integrative approach. In T. Angelovska & A. Hahn (eds.), *L3 syntactic transfer: Models, new developments and implications*. [BAP Vol. 5] (pp. 1-9). Amsterdam: John Benjamins.

Babatsouli, E., & Ingram, D. (2015). What bilingualism tells us about phonological acquisition. In R. H. Bahr & E. R. Silliman (eds.), *Routledge handbook of communication disorders* (pp. 173-182). Routledge: Taylor & Francis.

Bardel, C., & Falk Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, *23*, 459-84.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language.* Oxford, UK: Oxford University Press.

Falk, Y., & Bardel, C. (2010a). The study of the role of the background languages in third language acquisition. The state of the art. *International Review of Applied Linguistics*, *48*, 185-219.

Falk, Y., & Bardel, C. (2010b). Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research*, *27*, 59-82.

Hahn, A., & Angelovska, T. (2017). Input-practice-output: A method for teaching L3 English after L2 German with a focus on syntactic transfer. In T. Angelovska & A. Hahn (eds.), *L3 syntactic transfer: Models, new developments and implications*. [BAP Vol. 5] (pp. 299-319). Amsterdam: John Benjamins.

Haider, H. (2010). *The syntax of German*. Cambridge: CUP.

Kupisch, T. (2007). Determiners in bilingual German-Italian children: What they tell us about the relation between language influence and language dominance. *Bilingualism: Language and Cognition*, *10*, 57-78.

Llama, R., Cardoso, W., & Collins, L. (2010). The influence of language distance and language status on the acquisition of L3 phonology. *International Journal of Multilingualism*, *7*(1), 39-57.

Lloyd-Smith, A., Gyllstad, H., & Kupisch, T. (2016). Transfer into L3 English: Global accent in German-dominant heritage speakers of Turkish. *Linguistic Approaches to Bilingualism*, *7*(2), 131-162.

Rothman, J. (2013). Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance. In S. Baauw, F. Dirjkoninge & M. Pinto (eds.), *Romance languages and linguistic theory* (pp. 217-47). Amsterdam: Benjamins.

Sağın Şimşek, Ç. (2006). *Third language acquisition: Turkish-German bilingual students' acquisition of English word order in a German educational setting*. Münster/New York: Waxmann.

To cite this article: Angelovska, T. (2017). Phonologically induced syntactic transfer in L3 acquisition of English: Evidence from late sequential Russian-German bilinguals. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 21-27). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Third language production by early and late bilinguals

Rika Aoki
aokirika@mail.saitama-u.ac.jp

Saitama University

Abstract. The aim of this paper is to investigate whether knowledge of first language (L1) and/or second language (L2) has an impact on third language (L3) production of bilinguals. To answer this question, the present study conducted a production experiment and evaluated bilingual's L3 production. In the experiment, Mandarin-Japanese bilinguals, English monolinguals, Mandarin monolinguals, and Japanese monolinguals produced English initial plosives, and their voice onset time (VOT) values were measured. The bilinguals were divided into two groups depending on the age of arrival (AOA) in Japan to investigate the AOA effect in their L3 production. The raw data of the experiment showed that late bilinguals (AOA = 10-15 years old) and Japanese monolinguals generally produced shorter VOT than the other three groups. The results of ANOVA suggest that bilinguals produce L3 English plosives in a way that is not significantly different from English monolinguals. There were some exceptions in the results, however, according to the place of articulation and AOA: late bilinguals' VOT values were significantly different from those of English monolinguals in the production of alveolar plosives. Additionally, when bilinguals were compared with Mandarin and Japanese monolinguals, it was suggested that early bilinguals (AOA = 0-6 years old) were significantly distinct from their L2 Japanese monolinguals in labial and velar plosives, while late bilinguals were significantly different from their L1 Mandarin monolinguals in alveolar plosives. The results show that although bilinguals could somewhat adapt their L1 and/or L2 categories to L3, the range of VOT values associated with bilinguals' categories of English initial plosives were different according to their AOA in Japan.

Keywords: bilinguals, production, plosive, third language, voice onset time

Introduction

It has been claimed that people who have learned more than one language may be able to learn new languages more easily than those who have learned only one language. For example, bilingual students with knowledge of two orthographies showed significantly higher proficiency in the orthography/phonological decoding of their additional language, compared to students who had knowledge of only one orthographic type, suggesting that bilinguals tend to have higher orthographical and phonological ability (Abu-Rabia & Sanitsky, 2010). However, whether and how the knowledge of multiple languages facilitates the production of a newly acquired language (L3) and how the first language (L1) and/or second language (L2) of bilingual speakers affect L3 production remain unclear. To investigate these questions, the present study conducted a production experiment evaluating the production of bilinguals on L3 sound contrasts. More specifically, the experiment observed the production of initial stop voicing contrasts in English (e.g., /b, d, g/ vs. /p, t, k/) by bilinguals of Mandarin and Japanese who learned English as an L3, in terms of changes in voice onset time (VOT).

VOT is 'the time interval between the burst that marks release of the stop closure and the onset of quasi-periodicity that reflects laryngeal vibration' (Lisker & Abramson, 1964) and is a common indicator for aspiration in plosives. Generally, stop sounds are either of three modal ranges of VOT: lead (negative value), short-lag (positive but small value), and long-lag (positive and big value), and some languages use these VOT ranges differently to make phonological contrasts (Lisker & Abramson, 1964). For example, Mandarin stop contrast is based on the presence/absence of aspiration. Aspirated plosives, p^h, t^h, k^h, are produced as long-lag VOT averaging from 80 ms to 130 ms, and unaspirated plosives, p, t, k, are realized as short-lag VOT ranging between 5 ms and 25 ms (Shimizu, 1996). Meanwhile, Japanese has voiced and voiceless categories for stop contrast. Although their stop contrast does not rely on aspiration, Japanese voiceless plosives /p, t, k/ have short-lag VOT ranging from 30 to 66 ms, while voiced plosives /b, d, g/ have lead VOT between -89 ms and -75 ms (Shimizu, 1996). As for English, in which stop contrast is based on voicing, voiceless stops in a stressed syllable are pronounced as aspirated consonants although there is no aspirated consonant as an independent phoneme in English. English voiceless plosives /p, t, k/ exhibit long-lag VOT (68-88 ms), while voiced ones /b, d, g/ have short-lag VOT (7-22 ms) (Harada, 2007).

Thus, the distribution of VOT differs according to languages, and therefore, Mandarin-Japanese bilinguals are expected to have a different distribution of VOT from monolingual speakers of English, Mandarin, and Japanese. The present study investigated whether L3 VOT production of bilinguals is different from L1/L2/L3 monolinguals. In addition, the issue of whether bilinguals' age of arrival (AOA) in an L2 country affects their L3 VOT production was also explored.

Method

Participants

The experimental group of this study consisted of 16 native Mandarin speakers who immigrated to Japan before or around puberty, which means they were all Mandarin-Japanese bilinguals. All bilinguals resided in Japan and were undergraduate or high school students at the time of the experiment. The bilinguals were divided into two groups, in terms of their AOA in Japan. The participants who arrived in Japan between the age of 0 and 6 years old were categorized as early bilinguals (EB), while those who immigrated between 10 and 15 years old were referred to as late bilinguals (LB). Before the experiment, the bilinguals filled out a language background questionnaire about their linguistic background. A summary of the background information of our bilingual participants in this study is provided in Table 1.

Table 1. Background information on bilingual participants

ID	Group	Gender	Age	AOA	LOR
BL01	EB	F	20	0	20.1
BL02	EB	M	20	0	13.2
BL03	EB	M	20	0	20
BL04	EB	F	19	6	9
BL05	EB	M	18	4	14
BL06	EB	F	15	0	9
BL07	EB	F	19	0	18.75
BL08	EB	M	19	0	17
BL09	EB	M	19	2	13
BL10	EB	M	21	0	11
BL11	EB	F	20	0	17
BL12	LB	F	20	10	10.3
BL13	LB	F	24	12	12
BL14	LB	F	20	14	6.5
BL15	LB	F	18	15	2.5
BL16	LB	F	17	15	1.4

Note: Age = age at the time of the experiment (in years); AOA = age of arrival in Japan (in years); LOR = length of residence in Japan (in years)

As control groups, 8 native speakers of English, 10 native speakers of Mandarin, and 10 native speakers of Japanese took part in the experiment. English native speakers had scarce experience of nonnative languages, including Mandarin and Japanese. Mandarin and Japanese monolinguals had no or fairly limited experience of nonnative languages except English. They had received English language training at school in their mother country: they started to learn English as a foreign language around the age of 6 to 12. They had never lived in a country where English is overwhelmingly spoken.

Procedure

In the production experiment, the participants read aloud the English word list, which contained 12 CVC minimal pairs (i.e. 24 target words) with initial voiceless/voiced plosives followed by low vowels at a natural speech rate. The order of sentences was randomized for each participant. To make their production as natural as possible, only real words in each language were used, and they were read in the carrier sentence: *I say ___ to my friend*. Examples of the target words are shown in Table 2.

Table 2. English target words

Voiced IPA	**Spelling**	**Voiceless IPA**	**Spelling**
[baɪ]	bye	[pʰaɪ]	pie
[bæt]	bat	[pʰæt]	pat
[bæk]	back	[pʰæk]	pack
[bæn]	ban	[pʰæn]	pan
[daɪ]	die	[tʰaɪ]	tie
[daʊn]	down	[tʰaʊn]	town
[dæn]	Dan	[tʰæn]	tan
[daɪm]	dime	[tʰaɪm]	time
[gæp]	gap	[kʰæp]	cap
[gɑːɚd]	guard	[kʰɑːɚd]	card
[gæb]	gab	[kʰæb]	cab
[gʌm]	gum	[kʰʌm]	come
[baɪ]	bye	[pʰaɪ]	pie
[bæt]	bat	[pʰæt]	pat
[bæk]	back	[pʰæk]	pack

Recording took place in a soundproof recording booth at the University of Tokyo. The speech was recorded using a Marantz PMD660 digital recorder (24 bit, 48 kHz) and a SHURE SM58 microphone, a unidirectional dynamic vocal microphone. All sentences were repeated three times, which resulted in 24 × 3 = 72 speech samples for analysis.

Analysis

The recorded speech samples were analyzed acoustically to examine the VOT values of initial plosives in Mandarin words produced by the bilingual monolinguals. The VOT values were measured with the aid of wide-band spectrogram and waveform in Praat version 5.4.16 (Boersma & Weenink, 2015). Positive VOT was measured from the beginning of the burst to the onset of periodicity of the following vowel, while the length of negative VOT was determined by measuring the time between the beginning of periodic striations and the beginning of the burst.

Results

Figure 1 shows the mean English VOT values of five groups in the current study. As this table shows, bilinguals and Japanese monolinguals generally produced shorter VOT than the other three groups. With regard to voiced plosives, Japanese monolinguals were the only one group that produced lead VOT.

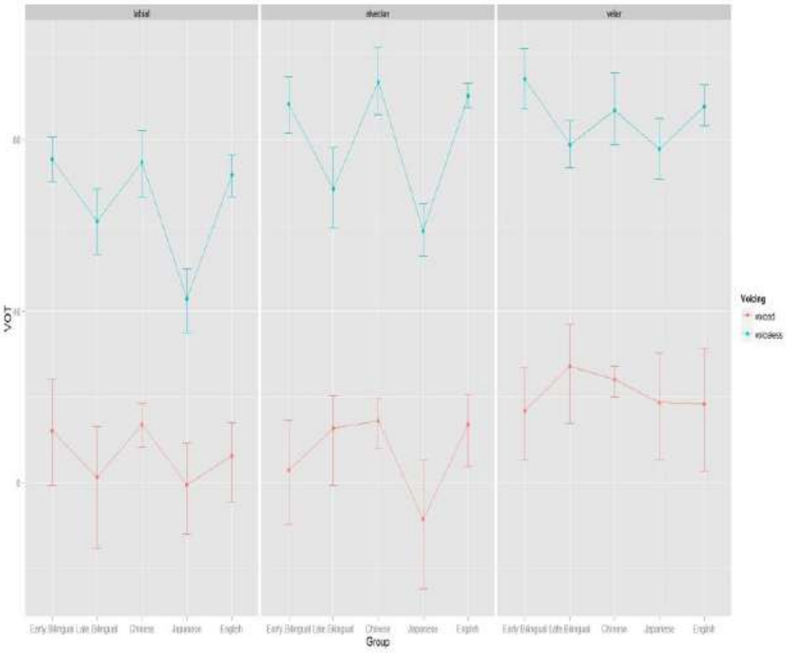


Figure 1. The mean VOT values of the voiceless and voiced plosives in English production

A mixed-effects ANOVA evaluating the effects of Group (early bilinguals, late bilinguals, Mandarin monolinguals, Japanese monolinguals, English monolinguals), Voicing (voiced, voiceless), and the interaction of Group × Participants, and the interaction of Group × Voicing was conducted for English VOT for labial, alveolar, and velar plosives. ANOVA conducted for labial plosives reported a significant effect of the interaction of Group × Voicing for a pair of early bilinguals and Japanese monolinguals ($F(1, 145) = 4.82$, $p = 0.03$), Mandarin and Japanese monolinguals ($F(1, 138) = 5.45$, $p = 0.021$), and Japanese and English monolinguals ($F(1, 124) = 7.06$, $p = 0.009$). As for alveolar plosives, the pairs of participant groups showing significant effects were the following: early bilinguals and late bilinguals ($F(1, 110) = 6.43$, $p = 0.013$), late bilinguals and Mandarin monolinguals ($F(1, 103) = 7.93$, $p = 0.006$), and late bilinguals and English monolinguals ($F(1, 89) = 6.06$, $p = 0.016$). The pair of Japanese and English monolinguals also had a marginally significant effect ($F(1, 145) = 3.36$, $p = 0.069$). ANOVA of velar plosives found that the three pairs of the participant groups differed significantly from each other: early bilinguals and Mandarin monolinguals ($F(1, 145) = 4.3$, $p = 0.04$), early bilinguals and Japanese monolinguals ($F(1, 145) = 4.81$, $p = 0.03$), and early bilinguals and late bilinguals ($F(1, 110) = 7.49$, $p = 0.007$).

Thus, it can be generally said that bilinguals produced L3 English plosives in a way that was not significantly different from English monolinguals, except for the production of alveolar plosives by late bilinguals. It is also suggested that bilinguals who arrived in Japan earlier produced English sounds in a significantly different way from bilinguals who arrived later in their childhood, at least for alveolar and velar plosives. Additionally, there was a tendency in which L3 English production of early bilinguals was significantly distinct from L2 Japanese monolinguals in labial and velar plosives, while that of late bilinguals was significantly different from L1 Mandarin monolinguals in alveolar plosives.

Discussion

From the present experiment, it was found that bilinguals generally did not show a significant difference in terms of English (L3) VOT from English monolinguals, except for alveolar stops produced by late bilinguals. This result implies that bilinguals might be sensitive to a wide range of VOTs, which allows them to establish new categories for L3 VOT, because Mandarin-Japanese bilinguals use a wide range of VOT values (from lead to long-lag) to process plosives in their L1 and

L2. This result differs from previous studies on L3 phonetic acquisition reporting influences from L1 (Llisterri & Poch-Olivé, 1987; Ringbom, 1987; Wrembel, 2012a; 2012b), L2 (Hammarberg & Hammarberg, 2005; Llama, Cardoso, & Collins, 2010; Wrembel, 2010), and L1/L2 combined effect (Sypiańska, 2013; Wrembel, 2014). The difference between the previous studies and the present study would lie in language distance: in previous studies, the fact that languages are genetically and also perceptually similar in terms of VOT would have had an influence on the results (e.g., Finish and Swedish, English and German), while the languages used in the present studies, namely Mandarin, Japanese and English, did not have both genetic and perceptual similarities.

With regard to the effect of AOA in Japan, the experiments showed that the range of VOT values associated with bilinguals' categories of English initial plosives were different according to their AOA: early bilinguals were significantly distinct from L2 monolinguals in labial and velar plosives, while late bilinguals were significantly different from L1 monolinguals in alveolar plosives. Here, it is indicated that early bilinguals tried to avoid being influenced by L2 Japanese, which is their dominant language, while late bilinguals tried to avoid being affected by Mandarin, their first acquired language. Thus, although early and late bilinguals generally succeeded in adjusting their categorical boundaries to produce English initial plosives, the range of VOT values associated with these categories was different according to their AOA in Japan.

Finally, let us compare this result with the predictions made by the Speech Learning Model (SLM) (Flege, 1995). In the assumption of SLM, a nonnative sound that is perceptually distinct from native sounds is more likely to be perceived and produced better through the establishment of a new phonetic category. However, if a learner does not perceive a phonetic difference between L1 and L2 sounds, he/she would be unable to forge a new category and possibly assimilate nonnative sounds into native categories. Here, SLM posits that if the L2 sound is close to but not identical to the L1 sound, the L2 sound cannot be produced native-likely. In the present study's context, although both English and Mandarin have long-lag VOT and both English and Japanese have short-lag VOT, Mandarin VOT is typically longer than English, while Japanese VOT is shorter than English. Therefore, SLM may predict that compared to English monolinguals, bilinguals would produce longer English voiceless VOT using their Mandarin long-lag VOT category and shorter English voiced VOT using their Japanese short-lag VOT category. Considering that the results of the present study showed that the VOT of bilinguals was not significantly different from that of monolinguals, the present study does not support the outcome of the prediction by SLM.

Conclusion

The results of the present study suggest that bilinguals generally showed no significant difference in L3 VOT production of L3 English initial plosives from English monolinguals. Thus, knowing two languages facilitated bilinguals' production of a wide range of VOT contrasts and, therefore, enabled them to produce L3 VOT accurately. The effect of AOA in the L2 country was also observed in the production of L3 alveolar and velar plosives. The present study has significant implications for studies of L3 phonetic acquisition, especially when the three languages of bilinguals are both genetically and perceptually distant. Additionally, the present study also showed that the predictions of L3 production inferred from SLM, which was originally a model of L2 speech processing, were incompatible with the results of L3 production experiments. Thus, it is suggested that SLM be modified for L3 production acquisition to predict a result more accurately. These results can be extended to the L3 phonetic acquisition of a bilingual in other settings, too.

References

Abu-Rabia, S., & Sanitsky, E. (2010). Advantages of bilinguals over monolinguals in learning a third language. *Journal of the National Association for Bilingual Education*, *33*, 173-199.

Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer* [Computer program]. Version 5.4.16, retrieved 17 August 2015 from http://www.praat.org/.

Flege, J. (1995). Two methods for training a novel second-language phonetic contrast. *Applied Psycholinguistics*, *16*, 425-442.

Hammarberg, B., & Hammarberg, B. (2005) Resetting the basis of articulation in the acquisition of new language: A third-language case study. In B. Hufeisen & R. Fouser, R. (eds.), *Introductory readings in L3* (pp. 11-18). Tubingen: Stauffenburg.

Harada, T. (2007). The production of voice onset time (VOT) by English-speaking children in a Japanese immersion program. *International Review of Applied Linguistics in Language Teaching (IRAL)*, *45*(4), 353-378.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops. *Word*, *20*, 384-422.

Llama, R., Cardoso, W., & Collins, L. (2010). The influence of language distance and language status on the acquisition of L3 phonology. *International Journal of Multilingualism*, *7*(1), 39-57.

Llisterri, J., & Poch-Olivé, D. (1987). Phonetic interference in bilingual's learning of a third language. *Proceedings of the XIth International Congress of Phonetic Sciences. Tallinn: Academy of Sciences of the Estonian SSR*, *5*, 134-147.

Ringbom, H. (1987). *The role of the mother tongue in foreign language learning*. Clevedon, UK: Multilingual Matters.

Shimizu, K. (1996). *A cross-language study of voicing contrasts of stop consonants in Asian languages*. Tokyo, Japan: Seibido Publishing Co.

Sypiańska, J. (2013). *Quantity and quality of language use and L1 attrition of Polish due to L2 Danish and L3 English*. Poznań: Adam Mickiewicz University. Unpublished PhD dissertation.

Wrembel, M. (2010). L2-accented speech in L3 production. *International Journal of Multilingualism*, *7*(1), 75-90.

Wrembel, M. (2012a). Foreign accent ratings in third language acquisition: The case of L3 French. In E. Waniek-Klimczak & L. Shockey (eds.), *Teaching and researching English accents in native and non-native speakers* (pp. 29-45). Berlin/Heidelberg: Springer Verlag.

Wrembel, M. (2012b). Foreign accentedness in third language acquisition: The case of L3 English. In J. Cabrelli Amaro, S. Flynn & J. Rothman (eds.), *Third language acquisition in adulthood* (pp. 281-309). Amsterdam: John Benjamins.

Wrembel M. (2014). VOT patterns in the acquisition of third language phonology. *Concordia Papers in Applied Linguistics (COPAL)*, *5*, 751-771.

To cite this article: Aoki, R. (2017). Third language production by early and late bilinguals. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 28-33). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

What does weak language mean in bilingual development? MLU vs. phonology

Elena Babatsouli
ebabatsouli@ismbs.eu

Institute of Monolingual and Bilingual Speech

Abstract. Studies concerning child bilingual acquisition compare phonological and/or morphological levels. It has been suggested (Bernardini & Schlyter, 2004) that weak is the language in which the child's mean length of utterance in words (MLUw) is one or more points lower than that in the dominant language. But does this language dominance in morphology guarantee dominance in phonology in the same language? This question is addressed in the present study by examining a bilingual child's English and Greek (two typologically different languages) over one month, at the age of 2;7. The female child acquired English in an exogenous environment (Greece) via her Greek mother's English speech. Her digitally recorded speech during daily mother/child interaction was transcribed orthographically and in IPA in a CLAN (MacWhinney, 2000) database. Because MLU cannot be compared between Greek and English, Greek being morphologically richer, MLUw is computed instead. The child's Greek MLUw is 3.36, and her English 2.01, indicating that Greek is the child's morphologically dominant language. By contrast, phonological levels match in the two languages, despite their typological differences (segmental inventories, phonotactics, etc.). Phonological level was assessed by the accuracy of singleton consonant production, of syllables, and of whole words including word length. Indicative results for English and Greek, respectively, are: whole word accuracy (44%, 45%), consonant-singletons words (52%, 52%), consonant-clusters words (11%, 8%). These results suggest that the morphologically dominant language boosts the bilingual child's phonological performance, though not the morphological performance in the weak language.

Keywords: bilingualism, child, development, consonant, speech performance, cross-linguistic measures

Introduction

Quantity of language exposure plays a key role in a child's speech progress during development; the more exposure the faster the progress. Speech progress is usually assessed in three ways: syntactically, morphologically, and phonologically. The present work addresses the question of whether, in a child's simultaneous bilingualism, more exposure in one language means that this language is stronger than the other language in terms of both morphology and phonology. The present study aims to complement results in child bilingual development which inform that, by around age 2;0, the two languages are differentiated (Genesee, Nicoladis, & Paradis, 1995; Nicoladis & Genesee, 1997), including their phonological systems (Keshavarz & Ingram, 2002), however interacting with each other and at times accelerating their acquisition, when compared to monolingual norms (Paradis & Genesee, 1996).

The present work examines whether the developing language with longer utterances in a child's bilingualism is also more advanced phonologically than the other language. The study combines a multi-level yardstick for quantitatively comparing bilingual performance across two languages that are morphologically different and vary in terms of word complexity, consonant inventories, and phonotactic rules, i.e. permitting different consonants and common consonants in different word positions. This is implemented by investigations on vocabulary, MLUw (or ALS), whole-word correctness and PMLU, individual consonant performance, and word complexity in terms of singletons and the number of syllables (e.g., Ingram & Dubasik, 2011).

Specifically, this is a case-study of a female child's bilingual speech in Greek/English, exploring aspects of child simultaneous bilingualism in an exogenous environment (Greece) for one of the languages (English), that enable re-assessment of the scope of universals and the extent to which developmental milestones in bilingualism are affected. Also, the present study addresses the shortage of studies in bilingual Greek/English (e.g., Antoniou, Best, Tyler, & Kroos, 2010), and a lack of studies examining the relationship between morphology and phonology in bilingual acquisition of Greek/Other-Language.

Quantitative measures for assessing morphological and phonological acquisition

The following outline introduces measures, used in the study, that facilitate morphological and phonological comparison of the bilingual child's languages.

Morphology

Utterance length has been a measure of child speech progress for almost one hundred years, named average length of sentence (ALS) (Nice, 1925), and renamed mean length of utterance in words (MLUw) following Brown's mean length of utterance in morphemes, MLU (Brown, 1973). Since one of the languages here is Greek, which is morphologically richer than the other language, English, MLU will not be an appropriate measure of language comparison, but MLUw will. Bernardini and Schlyter (2004) suggest that weak is the developing language in a child's bilingualism whose MLUw is at least one word less than that of the other language.

Phonology

Quantitative approaches to evaluating acquisition of phonology have been statistical in nature revealing the extent of acquisition of grammatical structures, measuring *correctness* of e.g., segments (e.g., Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997) or whole words (Ingram & Ingram, 2001). Production accuracy is determined as degree of correct use of a structure in obligatory environments with several criterions in use: 70% (e.g. Ingram, Christensen, Veach, & Webster, 1980), 75% (e.g., Diedrich & Bangert, 1980), 90% (Brown, 1973) correct production. In cross-sectional studies, correctness is viewed in terms of group performance (e.g. Smit, Hand, Freilinger, Bernthal, & Bird, 1990). The only quantitative approach to monolingual Greek phonology ascertaining norms at the 75% criterion of group performance is PAL (1995).

Schmitt, Howard, and Schmitt (1983) proposed the whole-word accuracy (WWA) formula during a single production of a given word; this is a measure that favorably complements existing measures of speech performance. With regard to singleton consonants performance, Shriberg *et al*.'s (1997) proportion of consonants correct (PCC) formula, computing the cumulative consonant correctness for each word type and then averaging them arithmetically, is widely accepted. Beyond these, assessment of phonological similarity among children has been viewed in terms of phonological mean length of utterance (PMLU) (Ingram & Ingram, 2001) which accounts for singletons performance in an utterance; unlike MLU (Brown, 1973) that counts all morphemes equally, PMLU doubles the count of correct-in-context consonant segments, because errors occur more often on targeted consonants (e.g., Ingram, 1981). The relation between produced and target words in child speech is accounted for by the phonological whole-word proximity (PWP) formula (Ingram & Ingram, 2001), which divides produced PMLU by the targeted PMLU. In a multi-dimensional phonological approach, Ingram and Dubasik (2011) propose nine measures that address whole-word, word position and prosodic aspects. A number of studies have used whole-word measures to quantify the level of phonological development in normal (e.g., Bunta, Fabiano-Smith, Goldstein, & Ingram, 2009) and phonologically impaired (Burrows & Goldstein, 2010) bilingual children.

Differences between Greek and English

As a fusional language, Greek is grammatically more complex than English: it is syllable-timed, with a simpler syllable structure ($C_{0\text{-}3}V_{0\text{-}1}$) and mostly longer words (bisyllabic and multisyllabic) than English (stress-timed), whose syllable structure is more complex ($C_{0\text{-}3}V_{0\text{-}4}$); the few monosyllabic words in Greek (Setatos, 1974) are function words. There is a larger proportion of consonants to vowels in stress-timed languages (Nespor, Peña, & Mehler, 2003), which also holds for English: English monosyllabic words are more complex, involving a larger variety of consonants and consonant clusters than Greek, reflecting the respective phonotactic differences of the two languages with regard to permissible word-final consonant singletons; with the exception of loans, only /n s/ are allowed in Greek word-finally.

With regard to their consonantal inventories, the two languages differ phonemically in that: Greek has two velar fricatives /x, ɣ/ while English has two palatoalveolars /ʃ, ʒ/ and glottal /h/; affricates in Greek are alveolar /ts, dz/ and in English palatoalveolar /tʃ, dʒ/; glides /j, w/ and the voiced velar nasal /ŋ/ are phonemic in English, but allophonic and/or dialectal in Greek; the rhotic in English is phonemically an approximant, /ɹ/, while in Greek it is a flap, /ɾ/. Among their common phonemes, allophonic processes (aspiration, palatalization, velarization, pre-nasalisation) and prosodic restrictions differentiate the two languages phonetically giving 11 allophones in Greek [ᶬb, ⁿd, ᵑg, c, ɟ, ŋ, ɱ, ɲ, ʎ, ç, ʝ] and 8 in English [pʰ, tʰ, kʰ, m̩, n̩, ɫ, l̩, ɹ̩]. A comprehensive summary of Greek phonology may be found in Mennen and Okalidou (2007). .

Method and participant

The child participant of the present study lives in Greece and was exposed to English from age 1;0 onwards through a single person's (her Greek mother's) English. The child attended a monolingual Greek daycare setting starting at age 2;0 and interacted with her mother (author) in English only at the remaining intervals (at home after daycare, during weekends and on vacations).

Data were elicited during these daily, routine speech interactions, they were digitally recorded, and also orthographically and IPA phonetically transcribed in a CLAN (MacWhinney, 2000) database. The data for the present study come from 41 CHAT files, each corresponding to a recording session, totaling in duration 3 hours over the month period from 2;7-2;8. There are 639 child utterances in Greek and 656 child utterances in English. Also, the child produced 307 English word types (1516 tokens) and 540 Greek word types (2374 tokens). This dataset is part of a larger database, where the child's Greek-English bilingualism is tracked longitudinally without interruption during 17 months of the child's bilingual development until age 4;0 (e.g., Babatsouli, 2016, 2017, 2018).

Results and discussion

First, the MLUw in the two languages will be calculated in order to determine whether they differ by more than one word and, thus, to determine whether one of the girl's languages is weak as compared to her other language. Repeated utterances and utterances with repeated words were excluded from the calculation. The results are shown in Table 1, where the number of utterances in each language is shown together with the percentage of utterances with a specific number of words.

Table 1. The bilingual girl's utterance length in words at age 2;7

L	U	MLUw	1w	2w	3w	4w	5w	6w	7w	8w	9w	10w	11w	12w
G	639	3.36	20%	21%	18%	16%	11%	8.1%	2.7%	1.3%	1.1%	1.3%	0.3%	0.2%
E	656	2.01	46%	24%	18%	7.5%	2.4%	0.9%	0.5%	0.2%	0%	0%	0%	0%

NB. L: Language, U: Utterances #, G: Greek, E: English, 1w, 2w: 1 word, 2 words, etc.

We see that English is the weak language, since its MLUw is 2.01, more than one word lower than the MLUw in Greek, which is 3.36. It is also observed that the child is capable of producing utterances that contain many more words than the average: as many as 12 words in Greek and 8 words in English, meaning that the child is capable of producing much longer utterances in both languages than the corresponding MLUw indicates. In any case, it is clear that English lags behind Greek in terms of utterance length. Compared to monolingual norms, the bilingual child's MLUw is slightly behind that of monolingual English children's (Brown, 1973), but ahead of monolingual Greek children's (compared to the children in Marinis, 2003).

The bilingual child's vocabulary at 2;7 consists of 307 English words and 540 Greek words, less than the vocabulary of monolingual English children at this age (e.g., Ingram, 1989) but more if the words in the vocabularies of the two languages are added, known to hold universally in bilingual children (Werker, Byers-Heinlein, & Fennell, 2009). Clearly, the child's utterance length as well as vocabulary size is found correlated to the amount of exposure and usage in each language.

Next, the bilingual girl's phonological level of acquisition of consonants is compared in her languages. The child's acquisition level of consonants is summarized in Table 2, where common consonants in the two languages are distinguished from language-specific ones.

Table 2. Acquisition level of consonants

% accuracy	common consonants	English specific	Greek specific
>50%	p b m f v t d s z l n j	ʃ ʒ tʃ dʒ	ʦ ɲ
<50%	θ ð k g	ɹ ɫ ŋ w h	ɾ ʎ, ç, ɟ x ç ɣ

We see that the acquisition level of consonants does not seem to be concomitant with the amount of exposure and usage in each language, but it is rather attributed to universal patterns of phonological acquisition (e.g., Bernhardt & Stemberger, 1998; PAL, 1995; Smit *et al*., 1990) and to the child's general preference of labials and coronals over palatals and velars in both languages. Compared to the acquisition level of monolingual English children (Smit *et al*., 1990), the bilingual girl here has acquired /p, b, m, s, z, n, l/ earlier, but acquires /k, g, w, h/ later. Her acquisition level in Greek is similar to that of monolingual Greek children (Mennen & Okalidou, 2007; PAL, 1995).

Accuracy in phonological performance permits more reliable evaluation of the acquisition of morphology, as in the case of English plural and 3rd person -s, and possessive <s> markers in both English and Greek, or in the acquisition of English possessives where simple and complex codas word-finally play a contrastive role (e.g. *my/mine, your/yours, her/hers*, etc.). In such contexts, coda deletion (e.g. *n*, *s*) may lead to ambiguity in the interpretation of the child's performance. However, the bilingual child of the present study acquires word-final *n*, *s* before the age of 3;0 and, therefore, her acquisition of possessives can be evaluated not only in terms of syntactic order but also of morphological accuracy. This has been studied by Babatsouli and Nicoladis (2016, 2017), who found that the child's Greek, with a higher MLUw and earlier acquisition of possessives, facilitated faster acquisition of the child's English possessives than the monolingual English norms, owing it to the more complex postnominal Greek possessive construction. Despite the morphological disparity and several surface differences between the possessives in the two languages, the bilingual child's developing grammar shows her underlying knowledge of their shared functional categories.

Attention is paid next to the cumulative performance of consonants in each language. For this purpose, Table 3 is presented where the percentage of consonants correct (PCC) is given separately over all consonants, over consonants common in the two languages, over language specific consonants, over singleton consonants, and over consonants clustering with other consonants.

Table 3. Percentage of consonants correct

language	overall	common consonants	language specific	singletons	clusters
Greek	62%	74%	10%	68%	39%
English	60%	68%	25%	64%	43%

It is seen that the child's overall consonant performance is at a similar level in both languages, despite phonological and phonotactic differences in the languages. Common consonants perform slightly better in Greek than in English. This fact is also reflected in the bilingual girl's words which contain on average 50% vowels in Greek, higher than the 43% vowel average in her English. Language-specific consonants perform better in English than in Greek, attributed to universal patterns of phonological acquisition and not to the amount of language exposure and usage. Singleton consonants perform slightly better in Greek than in English, attributed to the syllabicity of the Greek language. Consonants perform slightly better in English than in Greek when clustering with other consonants, attributed to language specific phonotactic rules. As in the case of individual consonant performance, it is concluded that cumulative consonant performance is similarly not influenced by the amount of exposure and usage, but it is attributed to universal tendencies and facilitative effects in the child's bilingualism, as the performance level is similar in both languages.

Lastly, the child's performance on whole word accuracy is compared in the two languages. To this end, Table 4 is presented next, where comparison is made separately over all words, over monosyllabic words, over multisyllabic words, over words with singleton consonants, and over words with consonant clusters. The last column shows the bilingual child's average phonological word proximity (PWP), as defined by Ingram (2002), whereby consonants are weighed twice as much as substituted consonants and produced vowels.

Whole word accuracy is at a similar level in both languages, as is also phonological word proximity again vouching for universal effects rather than amount of input and usage in each language. Monosyllabic words are performed better in Greek, mostly being function words without word-final consonant clusters as opposed to English, where the child's targeted words are mostly monosyllabic containing word-final consonant clusters. Multisyllabic words are performed similarly in the two languages as are words with singleton consonants and consonant clusters.

Table 4. Whole word accuracy

language	overall	monosyllabic	multisyllabic	singletons	clusters	PWP
Greek	45%	62%	34%	52%	8%	83%
English	44%	46%	35%	52%	11%	80%

Conclusion

It is concluded that in a Greek/English bilingual child's speech development, following language and phonological system differentiation, the larger the amount of exposure and usage in a language, the longer the utterances. In contrast, the child's morphologically weak language (English) does not lag behind the dominant language (Greek) phonologically. This is attributed to universal patterns of phonological acquisition, as well as to the amount of language exposure and usage cumulatively in the two languages and not in each language alone (facilitative effects from L1 to L2 in bilingual acquisition), resulting to positive interference accelerating the acquisition of phonology in the otherwise weak language.

References

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits nativelike stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics, 38*, 640-653.

Babatsouli, E. (2018). Acquired singleton fricatives and lateral in cluster development: A bilingual case. In E. Babatsouli & D. Ingram (eds.), *Phonology in protolanguage and interlanguage* (pp. 109-138). Sheffield, UK: Equinox Publishing.

Babatsouli, E. (2017). Bilingual development of *theta* in a child. *Poznan Studies in Contemporary Linguistics, 53*(2), 157-195

Babatsouli, E. (2016). Added syllable complexity in a child's developmental speech and clinical implications. *Clinical Linguistics and Phonetics, 30*(8), 628-648.

Babatsouli, E., & Nicoladis, E. (March 2016). Case study of a Greek-English bilingual child's possessives. Manuscript #3, Institute of Monolingual and Bilingual Speech, Chania, Greece.

Babatsouli, E., & Nicoladis, E. (June 2017). Case study of a Greek-English bilingual child's possessives: Input and crosslinguistic inferences. Manuscript #6, Institute of Monolingual and Bilingual Speech, Chania, Greece.

Bernardini, P., & Schlyter, S. (2004). Growing syntactic structure and code-mixing in the weaker language: The Ivy Hypothesis. *Bilingualism: Language and Cognition*, *7*(1), 49-69.
Bernhardt, B., & Stemberger, J. (1998). *Handbook of phonological development: From a nonlinear constraints-based perspective*. San Diego: Academic Press.
Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
Bunta, F., Fabiano-Smith, L., Goldstein, B. A., & Ingram, D. (2009). Phonological wholeword measures in three-year-old bilingual children and their age-matched monolingual peers. *Clinical Linguistics and Phonetics*, *23*, 156-175.
Burrows, L., & Goldstein, B. A. (2010). Whole word measures in bilingual children with speech sound disorders. *Clinical Linguistics and Phonetics*, *24*, 357-368.
Diedrich, W., & Bangert, J. (1980). *Articulation learning*. Houston, TX: College Hill Press.
Genesee, F., Nicoladis, E., & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, *22*, 611-631.
Ingram, D. (1981). *Procedures for the phonological analysis of children's language*. Baltimore, MD: University Park Press.
Ingram, D. (1989). *First language acquisition: Method, description and explanation.* Cambridge: Cambridge University Press.
Ingram, D. (2002). The measurement of whole-word production. *Journal of Child Language*, *29*, 713-733.
Ingram, D., & Dubasik, V. L. (2011). Multidimensional assessment of phonological similarity within and between children. *Clinical Linguistics and Phonetics*, *25*, 962-967.
Ingram, D., & Ingram K. (2001). A whole-word approach to phonological analysis and intervention. *Language, Speech and Hearing Services in Schools*, *32*, 271-283.
Ingram, D., Christensen, L., Veach, S., & Webster, B. (1980). The acquisition of word-initial fricatives and affricates in English by children between 2 and 6 years. In G. H. Yeni-Komshian, J. F. Kavanaugh & C. A. Ferguson, *Child phonology, Vol. I: Production* (pp. 169-192). New York, NY: Academic Press.
Keshavarz, M., & Ingram, D. (2002). The early phonological development of a Farsi-English bilingual child. *International Journal of Bilingualism*, *6*, 255-269.
MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum.
Marinis, T. (2003). *The acquisition of the DP in Modern Greek*. Amsterdam: John Benjamins.
Mennen, I., & Okalidou, A. (2007). Acquisition of Greek phonology: An overview. In S. McLeod (ed.), *The international guide to speech acquisition* (pp. 398-407). Clifton Park, NY: Thomson Delmar Learning.
Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, *2*, 203-229.
Nice, M. (1925). Length of sentence as a criterion of child progress in speech. *Journal of Educational Psychology*, *16*, 370-379.
Nicoladis, E., & Genesee, F. (1997). Language development in preschool bilingual children. *Journal of Speech-Language Pathology and Audiology*, *21*, 258-270.
PAL (Panhellenic Association of Logopaedics) (1995). *Assessment of phonetic and phonological development*. Athens: PAL. (in Greek)
Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or independent? *Studies in Second Language Acquisition*, *18*, 1-25.
Schmitt, L. S., Howard, B. H., & Schmitt, J. F. (1983). Conversational speech sampling in the assessment of articulation proficiency. *Language, Speech, and Hearing Services in Schools*, *14*, 210-214.
Setatos, M. (1974). *Phonology of Modern Greek*. Athens: Papazisis Publications (in Greek).
Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language and Hearing Research*, *40*, 708-722.
Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, *55*, 779-798.
Werker, J. F., Byers-Heinlein K., & Fennell, C. T. (2009). Bilingual beginnings to learning words. *Philosophical Transactions of the Royal Society of London B*, *364*, 3649-3663.

To cite this article: Babatsouli, E. (2017). What does weak language mean in bilingualism: MLU vs. Phonology. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 34-39). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

A measure for the proximity of consonant clusters in speech productions

Elena Babatsouli[1], Dimitrios Sotiropoulos[2]
ebabatsouli@ismbs.eu, dimsotirop@yahoo.com

[1]Institute of Monolingual and Bilingual Speech, [2]Technical University of Crete

Abstract. The proximity of consonant clusters in speech productions is measured in the literature by the proportion of correctly produced clusters to targeted clusters (e.g., Becker & Tessier, 2011; Smit, 1993). With this measure, all other cluster productions score 0%. However, several authors (e.g., Kirk & Demuth, 2005) support that clusters produced as two consonants in a sequence, independently of correctness, should be considered acquired. Therefore, a measure is needed that would quantitatively differentiate not only one-cluster-member productions from two-cluster-members (with or without vowel insertion) productions, but also cluster omission from cluster reduction. Within the framework of whole word proximity, Ingram (2002) proposed a measure for the proximity of produced consonants not only by accurate productions but also by substitutions, whereby substitutions weigh *half* as much as accurate productions. If this measure is applied to the proximity of consonant clusters, it will not differentiate reduced clusters, with one correct member, from clusters produced with both members substituted, nor clusters with vowel insertion from clusters produced as targeted. In the present study, a measure is proposed for the proximity of consonant clusters that differentiates all possible cluster productions, which are nine. The proposed measure is applied to typical and atypical speech productions in monolingual and bilingual children cross-linguistically, quantifying more realistically the proximity of consonant clusters than the correct/incorrect measure used thus far in the literature.

Keywords: measure, consonant clusters, acquisition, development, typical, atypical, monolingual, bilingual

Introduction

Consonant clusters, as compared to singleton consonants, make words more complex causing various errors in typical and atypical monolingual and bilingual child speech, as well as in adult impaired or second language speech. The effect of complexity on whole word production has been discussed in the literature both qualitatively (e.g., Stoel-Gammon, 2011) and quantitatively (e.g., Ingram, 2002; Babatsouli, Ingram, & Sotiropoulos, 2014; Kehoe, 2015b; Babatsouli *et al*., 2016; Babatsouli *et al*., 2017). Children have production problems with consonant clusters, normally until about age 4. Moreover, it is estimated that about 5% of six-year-old children have speech sound disorders (e.g., Broomfield & Dodd, 2004), delaying the acquisition of consonant clusters. These children require intervention targeting clusters that consist of members that have either been acquired as singletons or that have not been acquired as singletons. In fact, it is expected that treatment on clusters will cause improvements on singletons in the child's sound system (e.g., Barlow, 2004).

Error patterns in the production of consonant clusters have been well documented both in children with typical development (e.g., Greenlee, 1974; Smit, 1993), as well as in children with speech sound disorders (e.g., Chin & Dinnsen, 1992; Ingram, 1972; McLeod, van Doom, & Reed, 1997). The most common error is cluster reduction, whereby one cluster member is deleted. Wyllie-Smith, McLeod, and Ball (2006) report for typically developing and speech-impaired children that when clusters reduce to a target consonant, it is the least sonorant cluster member while, when they reduce to a non-target consonant, the sonority hypothesis is not adhered to. There are three main stages in cluster development: reduction, two- member production with one or both cluster members substituted, and acquisition (Greenlee, 1974; Smit, 1993). For typically developing children, these stages are age-dependent and successive, even though there is overlapping between them, with the earlier stage being the reduction stage. Not all consonant clusters go through the development stages at the same age, even for a single child. Some clusters, usually the ones with cluster-members whose sonority distance is small, have been reported as harder to produce than others, for monolingual English children (e.g., Barlow, 2004). For typical monolingual English children, Smit (1993) found the following order of acquisition (75% accuracy) with increasing age: stop + w, l clusters, r clusters, s clusters, sl and θr.

Several authors, including Kirk and Demuth (2005), consider that the main difficulty in producing a consonant cluster is to produce consonants adjacent to each other, independently of whether they are produced as targeted. In other words, clusters that are produced as substitutions should be considered acquired. This claim is based on the fact that before correct cluster production, children first reduce clusters and, later on, they often insert a vowel between cluster-members. To date, there is no study in the literature that quantitatively reflects this when calculating percentages of clusters acquired. Calculations are performed based on the percentage of correctly produced clusters to targeted clusters. In these calculations, the percentage of accuracy of substituted clusters is 0%, as is the accuracy percentage of clusters with a vowel inserted between cluster-members, as well as the accuracy percentage of reduced clusters (e.g., Becker & Tessier, 2011; Smit, 1993) and, thus, they cannot be differentiated quantitatively from each other.

The following research question, thus, arises: how should the proximity of clusters be measured in order to quantitatively capture different stages in cluster development and, also, consider substituted clusters as acquired? In other words, what kind of measure can differentiate quantitatively all nine possible productions of a two-member consonant cluster? The nine possible productions are: omission/reduction to one of the target consonants that is substituted, reduction to a target consonant, two substituted members with an added vowel between them, one correct and one substituted member with an added vowel between them, two correct members with an added vowel between them, two substituted members in sequence as in the target cluster, one substituted member and one correct member in sequence as in the target cluster, two correct members in sequence as in the target cluster. The question is addressed in the present paper by proposing a measure that achieves this.

Before going into the main part of the paper, it may be of interest to examine whether the PWP measure proposed by Ingram and Ingram (2001) and Ingram (2002) for whole word phonological proximity can be applied to the proximity of consonant clusters and, if so, what the result will be. They proposed that deletions be given zero points, substituted consonants and produced vowels one point, and correctly produced consonants two points. Further, target consonants are given two points and target vowels are given one point. If such a measure is applied to the proximity of consonant clusters, it does not differentiate clusters reduced to one target consonant from clusters with both members substituted, both having 50% proximity. Furthermore, it does not differentiate consonant clusters produced with a vowel inserted between the cluster members from consonant clusters produced with no added vowel, both having the same proximity, 50% or 100% depending on whether the target consonants are substituted or are correctly produced.

The proposed measure of cluster proximity

The proposed measure for cluster proximity assigns zero points to a deleted cluster member, one point to a substituted member, two points to a correctly produced member, one extra point when two cluster members are produced with a vowel inserted between them, and four extra points when two cluster members are produced without an added vowel between them. Cluster proximity is defined as the ratio of the total points of the produced cluster to the total points of the target cluster, which are 8. Then, the measure results in the proximity scale for all possible productions of two-member consonant clusters, shown in Table 1.

Based on the correct/incorrect measure used in the literature for consonant clusters (e.g., Becker & Tessier, 2011; Smit, 1993), all the above cluster productions score 0%, except the two-member correct productions which the measure scores as 100%. Furthermore, if Ingram's (2002) scoring for measuring the proximity of whole-word productions is applied to the proximity of consonant cluster productions, it yields: DD: 0%; SD, DS: 25%; CD, DC: 50%; SVS: 50%; CVS, SVC: 75%; CVC: 100%; SS: 50%; CS, SC: 75%, CC: 100%. In such a scale, reduced clusters are not differentiated from two-member productions which, in turn, are not differentiated from vowel insertion between the two cluster members.

Table 1. The proposed measure for cluster proximity

Process	**omission/reduction**			**vowel insertion**			**two-member production**		
production	**DD**	**SD, DS**	**CD, DC**	**SVS**	**CVS, SVC**	**CVC**	**SS**	**CS, SC**	**CC**
proximity	**0%**	**12.5%**	**25%**	**37.5%**	**50%**	**62.5%**	**75%**	**87.5%**	**100%**

NB. D: Deleted, S: Substituted, C: Correct, V: Vowel Inserted

Examples of cluster productions by monolingual and bilingual children with typical and atypical development that correspond to the above categories of the proximity measure follow.

Omission (DD):

Normal speech
bilingual German-Spanish (first 50 words): *plato* [ʔato] (Kehoe, 2015a)
monolingual English (2;6): *fly* [ʋai], *different* [dipi:] (Smith, 1973)
Disordered/Delayed speech
monolingual English (3;0): *finger* [i.ʌ], *jumping* [ʌ.ə], *orange* [ɔ.ʌ], *sleeping* [i.ɪn] (Ingram, 2002)
monolingual English (4;8): *cry* [ai], *play* [eɩ], *spoon* [ʊn], *star* [a], *sky* [ai] (Chin & Dinnsen, 1992)

Reduction to a non-target (SD, DS):

Normal speech
monolingual Spanish (first 50 words): *otro* [œdɔ] 'other' (Kehoe, 2015a)
bilingual German-Spanish (first 50 words): *drei* [kaɪ] 'three', *manzana* [pada] 'apple' (Kehoe, 2015a)
bilingual Greek-English (2;10): scepaso [tepaso] 'to cover'; *swing* [fɪn] (Yavaş & Babatsouli, 2016)
Disordered/Delayed speech
monolingual English (3;0): *flag* [baɪ], *butterfly* [bʌbaɪ] (Ingram, 2002)
monolingual English (3;11): *play* [reɩ] (Chin & Dinnsen, 1992)

Reduction to a target (AD, DA):

Normal speech
monolingual English (2;4): *green* [gi:n] (Smith, 1973)
bilingual Greek-English (2;10): sfiga [figa] 'bee', sxolio [solio] 'school' (Yavaş & Babatsouli, 2016)
Disordered/Delayed speech
monolingual English (3;0): *blu* [bu] (Ingram, 2002)
trilingual English-German-French (4;0): *grand* [gã], *drin* [dɪn] 'in it' (Kehoe, 2015b)

Vowel Insertion (CVC):

Normal speech
monolingual English (2;6): *glue* [gu:l] (Smith, 1973)
Disordered/Delayed speech
monolingual English (3;11): *basket* [səkə], *desk* [sək] (Ingram, 1972)

2-Member Production (SS):

Normal speech
monolingual English (2;6): *green* [kli:n] (Smith, 1973)
bilingual Greek-English (2;10): *sleep* [ʦi:p], *swim* [ʦim] (Yavaş & Babatsouli, 2016)
Disordered/Delayed speech
monolingual English (3;11): *crib* [flɪb] (Ingram, 1972)
monolingual English (3;11): *clean* [gjin] (Chin & Dinnsen, 1992)

2-Member Production (SC, CS):

Normal speech
monolingual English (2;6): *fly* [vlaɪ], *soft* [wɔpt) (Smith, 1973)
bilingual Greek-English (2;10): sxolio [stoli] 'school'; *sleep* [tli:p] (Yavaş & Babatsouli, 2016)
Disordered/Delayed speech
monolingual English (3;11): *cloth* [fləf] (Ingram, 1972)

2-Member Production (CC):

Normal speech
monolingual English (3;3): *fly*, *flower*, *soft* (Smith, 1973)
bilingual Greek-English (2;10): *small* [smɒl], *snail* [sneɪɫ] (Yavaş & Babatsouli, 2016)
Disordered/Delayed speech
monolingual English (3;11): *spoon*, *swim*, *sleeping*, *smell*, *sky*, *stove* (Barlow, 2001)

Conclusion

A measure for the proximity of consonant clusters was proposed that distinguishes quantitatively all possible cluster productions. The validity of the measure and its advantages over the correct/incorrect measure used in the literature were shown by utilizing cross-linguistic, monolingual and bilingual, typical and atypical developmental child speech data available in the literature. It is aimed that the proposed measure will be used for assessing productions of consonant clusters in typical and atypical child speech, as well as in impaired adult and second language speech cross-linguistically, and guide intervention whenever necessary.

Acknowledgements

Preliminary results of this work were presented at the *International Child Phonology Conference 2016* in Flagstaff, Arizona. We wish to thank Sharon Inkelas of the University of California, Berkeley and David Ingram of Arizona State University for their suggestion to include vowel insertion in our proposed measure.

References

Babatsouli, E., Ingram, D., & Sotiropoulos, D. A. (2014). Phonological word proximity in child speech development. *Chaotic Modeling & Simulation*, *4*(3), 295-313.

Babatsouli, E., Ingram, D., & Sotiropoulos, D. (2016). Entropy as a measure of mixedupness of realizations in child speech. *Poznan Studies in Contemporary Linguistics*, *52*(4), 605-627.

Babatsouli, E., Ingram D., & Sotiropoulos, D. (2017). On the weight of segments in phonological word proximity. In E. Babatsouli, D. Ingram & N. Müller (eds.), *Crosslinguistic encounters in language acquisition: Typical and atypical development* (pp. 51-70). Bristol, UK: Multilingual Matters.

Barlow, J. A. (2001). The structure of /s/-sequences: Evidence from a disordered system. *Journal of Child Language*, *28*, 291-324,

Barlow, J. A. (2004). Consonant clusters in phonological acquisition: Applications to assessment and treatment. *California Speech Hearing Association Magazine*, *34*, 10-13.

Becker, M., & Tessier, A.-M. (2011). Trajectories of faithfulness in child-specific phonology. *Phonology, 28*, 163-196.

Broomfield, J., & Dodd B. (2004). Children with speech and language disability: Caseload characteristics. *International Journal of Language and Communication Disorders*, *39*(3), 303-24.

Chin, S. B., & Dinnsen, D. A. (1992). Consonant clusters in disordered speech: Constraints and correspondence patterns. *Journal of Child Language, 19*, 259-285.

Greenlee, M. (1974). Interacting processes in the child's acquisition of stop-liquid clusters. *Papers and Reports on Child Language Development* (Stanford University), *7,* 85-100.

Ingram, D. (1972). Phonological analysis of a developmentally aphasic child. Institute of Childhood Aphasia. Stanford University School of Medicine.

Ingram, D. (2002). The measurement of whole-word production. *Journal of Child Language*, *29*, 713-733.

Ingram, D., & Ingram, K. (2001). A whole-word approach to phonological analysis and intervention. *Language, Speech and Hearing Services in Schools*, *22*, 271-283.

Kehoe, M. (2015a). Lexical-phonological interactions in bilingual children. *First Language*, *35*(2), 93-125.

Kehoe, M. (2015b). Acquisition of velars: a whole-word approach. *Clinical Linguistics and Phonetics*, *29*(12), 873-908.

Kirk C., & Demuth, K. (2005). Asymmetries in the acquisition of word-initial and word-final consonant clusters. *Journal of Child Language*, *32*(4), 709-734.

McLeod, S., van Doorn, J., & Reed, V. A. (1997). Realizations of consonant clusters by children with phonological impairment. *Clinical Linguistics and Phonetics*, *11*(2), 85-113.

Smit, A. B. (1993). Phonological error distributions in the Iowa-Nebraska Articulation Norms Project: Word-initial consonant clusters. *Journal of Speech and Hearing Research*, *36*, 931-947

Smith, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge, UK: Cambridge University Press.

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language, 38*, 1-34.

Wyllie-Smith, L., McLeod, S., & Ball, M. J. (2006). Typically developing and speech-impaired children's adherence to the sonority hypothesis. *Clinical Linguistics and Phonetics*, *20*(4), 271-291.

Yavaş, M., & Babatsouli, E. (2016). Acquisition of /s/-clusters in a Greek-English bilingual child: Sonority or OCP? In M. J. Ball (ed.), *Challenging sonority: Cross-linguistic evidence from normal and disordered language* (pp. 337-354). Sheffield, UK: Equinox Publishing.

To cite this article: Babatsouli, E,. & Sotiropoulos, D. (2017). A measure for the proximity of consonant clusters in speech productions. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 40-44). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Self-organisation in phonological development: *Templates* in Brazilian and European Portuguese

Maria de Fátima de Almeida Baia[1], Susana Correia[2]
mariadefatimabaia@uesb.edu.br, susanacorreia@fcsh.unl.pt

[1]Universidade Estadual do Sudoeste da Bahia, [2]Universidade Nova de Lisboa

Abstract. This study investigates the emergence of early phonological *templates*, i.e. systematic shapes that facilitate the expansion of the lexicon, in the linguistic development of six children, three acquiring Brazilian Portuguese (BP) and three acquiring the European variety (EP). We follow the dynamic perspective of phonological development, which understands language development as a process of evolution characterised by variability. According to this view, although there is instability in development there are adjustments in the system due to the self-organization principle, which can be understood as a spontaneous pattern formation. The whole-word phonology/templates model is adopted in the analysis of early phonological patterns. The model claims that the organising principle in early phonological development is the whole word, not features or segments. These early word patterns are understood as templates. A preliminary longitudinal study was conducted to investigate early templates in BP and EP. We observed variability between the subjects regarding *template* types. In addition, we observed that instances of use and disuse of templates vary from child to child. The preliminary results suggest that there are templates operating during variability in phonological development. Also, the manifestation of a template characterises instability in the development. In instances of instability, templates are formed due to the principle of self-organisation, namely the spontaneous formation of patterns. The system organises itself due to its inherent ability to find patterns from some type of interaction.

Keywords: Brazilian and European Portuguese, dynamic systems, self-organization, templates

Introduction

This study investigates the emergence of early phonological *templates* in the phonological development of three children acquiring Brazilian Portuguese (BP) and three children acquiring the European variety (EP).

Within the dynamic systems theory in the study of child language, variety, flexibility and asynchrony tend to occur in the developmental process (Thelen & Smith, 1994). In this perspective, there is instability in development as there are adjustments in the system due to a self-organisation principle. Dynamic systems theory stresses the continuity between the development of phonological structure and the development of all other structures in nature (Szreder, 2012:14). Self-organisation thus appears as ‘order emerging without hierarchical pre-planning, based on the structural and functional capacities of the system’ (Davis & Bedore, 2013:134) and as a result of spontaneous pattern formation (Kelso, 1995). In an emergentist view, self-organisation does not require a blueprint or cookbook (Davis & Bedore, 2013:156).

Within the perspective of whole word/templatic phonology, the whole word is understood as the organising principle in early phonological development, not features or segments (Vihman & Croft, 2007; Vihman & Velleman, 2002). Many unusual phonological substitutions tend to occur due to ‘pattern force’ (Macken, 1979). In this perspective, some early word patterns are understood as templates, i.e. systematic shapes that facilitate the expansion of the lexicon.

The whole word approach consists of an attempt to understand children’s phonological development in itself and on its own terms (Ferguson & Farwell, 1975). Word templates are seen as child-specific word patterns and their effect is to make a lot of the child’s words sound similar to each other (Keren-Portnoy, Vihman, DePaolis, Bidgood, & McGillion, 2011). They are the child’s responses to challenging target forms. Difficulties for the child emerge due to limitations/idiosyncrasies on articulation, articulatory planning, memory or biases/preferences. Therefore, idiosyncrasies give rise to individual differences in production. Moreover, individual templates may be similar to the target word (*selected forms*) or be distorted forms of the target word (*adapted forms*). Some examples of 'bizarre' forms in the acquisition of English as L1 have been reported previously (Waterson, 1971:179):

Child Production	**Target**
[ˈɲẽ.ɲẽ]/[ˈɲi.ɲi]	*finger*
[ˈɲẽ.ɲẽ]	*window*
[ˈɲã.ɲã]	*another*
[ˈɲã.ɲø]	*Randall*

In Brazilian Portuguese, some examples of these 'bizarre' forms, emerging often as reduplicated syllables, have also been found. Baia (2008) names these 'bizarre', adapted forms produced by a child 'lexical creations', following Secco (1994). The same is observed in the acquisition of European Portuguese (Correia, 2010). Therefore, adapted and selected forms can, in fact, be identified in the early speech of children speaking two varieties of Portuguese, BP and EP:

Table 1. Child words in BP and EP acquisition

Brazilian Portuguese (Baia 2010, 2013)				**European Portuguese (**Baia & Correia, 2016; Correia 2010)			
Age	*Child production*	*Target*	*Gloss*	*Age*	*Child production*	*Target*	*Gloss*
1;1	[ba.ˈba] [a. ‘bo]	*bola* [ˈbɔla]	'ball'	1;1	[ˈpəˈbə] [ˈbaˈbɐ]	*bebé* [bɛˈbɛ]	'baby'
1;3	[ba.'ba]			1;2	[aβɐˈβɛ] [ɐˈβɛβe]		
1;5	[ba.'ba] [bɔ.'bɔ]			1;4	[ɐbɐˈbe] [ɛbeˈbe]		
1;7	[bɔ.'bɔ] [ˈbɔ.ja]			1;7	[beˈbe]		
1;1	[ka.'ka] [ka]	*Karine* [kaˈɾini]	'name of the child's sister'	1;1	[ˈɐ̃ːw]	*não* [nɐ̃w̃]	'no'
1;3	[ka.'ka]			1;2	[nɐˈnɐ̃w̃]		
1;5	[ka.'ka]			1;4	[na] [nɐ̃w]		
1;7	[ka.ˈka] [ka.ˈi]			1;7	[nɐ̃w̃]		

The productions illustrated show that the speech of children speaking BP and EP is mainly characterized by an initial iambic pattern, although Portuguese words have stress mainly in the penultimate syllable. The authors observe similar deletion and filler sound insertion strategies in the speech of Brazilian and Portuguese children. However, different word shapes are found, both within the same child and between children, across varieties.

In this paper, we hypothesize that inter- and intravariability found in the comparison of templates in different Brazilian children (Baia, 2013) will be similar to the variability found in the comparison between Brazilian and Portuguese children. That is, we expect to find in two varieties of the same language a variable developmental path.

Method

This study investigates the emergence of early phonological templates in the linguistic development of children acquiring Brazilian and European Portuguese. To conduct this study, we considered a sample of longitudinal spontaneous speech from 6 children (3 acquiring BP + 3 acquiring EP), between 0;9 and 2;0. For BP, we used the *corpus A aquisição do ritmo em Português Brasileiro – Processos de Ancoragem* (Santos, 2005). For EP, we used the 'CCF' *corpus*, from the *Acquisition of European Portuguese* databank (available in http://phonbank.talkbank.org/access/Romance/Portuguese/CCF.html). Children's speech was collected monthly in observational sessions of 30-45 minutes from both *corpora*. For the analysis, orthographic and IPA phonetic transcription were considered (with 90% inter-judge reliability; dubious 10% were excluded). To distinguish reduplicated words from babbling, we adopted the criteria proposed by Vihman and McCune (1994), i.e. the context (mother identification) and the type of vocalization (correspondence with the target word) disambiguated the child's utterance as *word* or *babbling*. Dubious and unintelligible transcriptions and productions were disregarded from the analysis. For the template analysis, a phonological structure had to occur at least in 40% of the total number of tokens in each session to be considered a template. Table 2 summarizes the data analysed.

Table 2. Data from Brazilian and Portuguese children analysed

Child	Age	Total # of productions
BP1 (male/Sao Paulo) Brazilian	0;9 – 2;0 16 sessions/months	Words: 1975 tokens
BP2 (male/ Sao Paulo) Brazilian	0;9 – 2;0 16 sessions/months	Words: 697 tokens
BP3 (male/ Sao Paulo) Brazilian	0;10 – 2;0 15 sessions/months	Words: 939 tokens
EP1 (female/ Lisbon) Portuguese	0;10 – 1;8 10 sessions/months	Words: 394 tokens
EP2 (female/ Lisbon) Portuguese	0;11 – 1;3 4 sessions/months	Words: 557 tokens
EP3 (male/ Lisbon) Portuguese	1;3 – 1;7 4 sessions/months	Words: 492 tokens

Results

Brazilian Portuguese

We observed variability between the subjects, regarding the type of template. In addition, we observed that instances of use and disuse of templates vary from child to child.

The following tables show how each template (T) was distributed along the sessions for the three Brazilian children observed. The first transcription correponds to the child's utterance, whereas the target appears in the second transcription.

We observe that the first child, BP1, uses templates for a reduced period of time when compared to his peers. BP1 uses templates until 1;4, whereas BP2 and BP3 use phonological patterns until 1;8 and 1;10, respectively. Full ($C_1V_1.ˈC_1V_1$) and partial ($C_1V_1.ˈC_1V_2$) reduplications are present all along the observational period, often simultaneously. Although the three children use reduplication as a production strategy, BP2 also uses V.'CV and 'V.CV words, BP1 and BP3 also use CV words as templates. Templates are in use and disuse as development proceeds. As the tables show, all Brazilian children whose data were analysed used selected and adapted templates, and that is independent of target word size or shape.

Table 3. BP1’s templates (BP)

	0;10	0;11	1;0	1;1	1;2	1;3	1;4
T	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2			CV	CV	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2

BP1: **i. reduplication (C_1V_1.ˈC_1V_1 and C_1V_1.ˈC_1V_2)**

Child form / *Target form*
[neˈne] *nenê* / [neˈne] ‘baby’ (selected)
[poˈpo] *vovô* / [voˈvo] ‘grandfather’ (selected)
[kaˈka] *galinha* / [gaˈlĩɲɐ] ‘hen’ (adapted)
[taˈta] *tchau* / [tʃaʊ] ‘bye’ (adapted)

ii. CV

Child form / *Target form*
[pɛ] *pé* / [pɛ] ‘foot’ (selected)
[la] *lá* / [la] ‘there’ (selected)
[da] *dado* / [ˈdadʊ] ‘dice’ (adapted)
[fo] *flor* / [floɾ] ‘flower’ (adapted)

Table 4. BP2’s templates (BP)

	0;9	0;10	0;11	1;3	1;4	1;5	1;6	1;8
T	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	V_1.ˈCV_1 V_1.ˈCV_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	ˈV_1. CV_1 ˈV_1.CV_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	C_1V_1.ˈC_1V_2

BP2: **i. reduplication (C_1V_1.ˈC_1V_1 and C_1V_1.ˈC_1V_2)**
Child form / *Target form*
[kʰo'kʰo] *cocô*/ [koˈko] ‘poo’ (selected)
[pa'pa] *sapato*/ [saˈpatʊ] ‘shoe’ (adapted)

iii. V.ˈCV
Child form / *Target form*
[a'ki] *aqui* / [aˈki] ‘here’ (selected)
[o'kʰo] *cocô* / [koˈko] ‘poo’ (adapted)

ii. ˈV.CV
Child form / *Target form*
['esɪ] *esse* / [ˈesɪ] ‘this’ (selected) ['opa] *roupa* / [ˈhopɐ] ‘clothes’ (adapted)

Table 5. BP3’s templates (BP)

	0;11	1;0	1;1	1;2	1;3	1;4	1;5	1;6
T	CV		C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	ˈV_1. CV_1 ˈV_1.CV_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2

	1;7	1;8	1;9	1;10	1;11	2;0
T	Cv_1V_1.ˈCv_1V_1 Cv_1V_1.ˈCv_1V_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2	C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2		

BP3: **i. reduplication (C_1V_1.ˈC_1V_1 and C_1V_1.ˈC_1V_2)**
Child form / *Target form*
[maˈmã] *mamãe* / [maˈmãĩ] ‘mother’ (selected)
[poˈpo] *alô* / [aˈlo] ‘hello’ (adapted)

iii. CV
Child form / *Target form*
[da] *dá* / [da] ‘give me’ (selected)
[kʰa] *caca* / ['kakɐ] ‘poo/dirty thing’ (adapted)

ii. ˈV.CV
Child form / *Target form*
['uma] *uma* / [ˈũmɐ] ‘one’ (selected)
['ɔkʰi] *óculos* / [ˈɔkulʊs] ‘glasses’ (adapted)

European Portuguese

As in BP data, we observed variability between the subjects regarding the type of template, as well as the instances of use and disuse of *templates.*

The following tables show how each template (T) was distributed along the sessions for the three Portuguese children observed:

Table 6. EP1's templates (EP)

	1;0	1;1	1;2	1;3	1;4	1;5	1;6	1;7	1;8
T	V_1.ˈCV_1 V_1.ˈCV_2		C_1V_1.ˈC_1V_1 C_1V_1.ˈC_1V_2			V_1.ˈCV_1 V_1.ˈCV_2			

EP1: **i. V.CV**
Child form / *Target form*
[ɐ.ˈla] *olá* / [ɔˈla] 'hello' (selected)
[a.ˈβe] *bebé* / [bɛˈbɛ] 'baby' (adapted)
[ɐˈpe] *pinguim* / [pĩˈgʷĩ] 'penguin' (adapt)
[ɐˈkɐ] *cão* [kɐ̃w̃] / 'dog' (adapted)

ii. reduplication (C_1V_1.ˈC_1V_1 and C_1V_1.ˈC_1V_2)
Child form / *Target form*
[βɨˈβe] *bebé* / [bɛˈbɛ] 'baby' (selected)
[tɐ.ˈta] *está* / [ʃˈta] 'it is' (adapted)
[paˈpa] *papa* / [pɐˈpa] 'daddy' (selected)

Table 7. EP2's templates (EP)

	0;11	0;12	1;1	1;2	1;3
T		CV	ˈC_1V_1.ˈC_1V_1		CV

EP2: **i. CV**
Child form / *Target form*
[mɐ̃] *mama* / [mɐˈmɐ̃] 'mommy' (adapted)
[ɲːe] *Inês* / [iˈneʃ] 'name' (adapted)
[tɔ] *toma* / [ˈtɔmɐ] 'take it' (adapted)
[da] *dá* / [da] 'give me' (selected)
[ʀɘ] *carro* / [ˈkaʀu] 'car' (adapted)

ii. reduplication (ˈC_1V_1.ˈC_1V_1)
Child form / *Target form*
[ˈmɐ.ˈmɐ̃] mamã / [mɐˈmɐ̃] *mother* (selected)
[ˈɲɐ.ˈɲɐ] não / [ˈnɐ̃w̃] *no* (adapted)

Table 8. EP3's templates (EP)

	1;3	1;4	1;5	1;6	1;7	1;8	1;9
T	CV			C_1V_1.ˈC_1V_1 CV		C_1V_1.ˈC_1V_2 CV	

EP3: **i. CV**
Child form / *Target form*
[pa] papá / [pɐˈpa] *daddy* (adapted)
[dɑ] dá / [da] *give me* (selected)
[βo] bola / [ˈbɔlɐ] *ball* (adapted)

ii. reduplication (C_1V_1.ˈC_1V_1 and C_1V_1.ˈC_1V_2)
Child form / *Target form*
[mɐˈmɐ̃] mama / [mɐˈmɐ̃] *mommy* (selected)
[dɐˈdɐ] dá / [da] *give me* (adapted)

Like Brazilian children, Portuguese children use reduplication as a production strategy. However, differences are observed between the children speaking the two varieties, as Portuguese children seem to use *templates* to a much lesser extent and for a smaller period of time. EP1 and EP3 use *templates* until 1;5 and 1;8, respectively, but EP2 abandons phonological patterns very early on (1;3). In Portuguese children, full and partial reduplication may co-occur (as in C), or they can alternate (as in

EP3). In addition to reduplication, which is a production strategy common to the Portuguese children observed, V.'CV (in EP1's case) and CV (in EP2 and EP3's case) are also available strategies.

As for Brazilian children, Portuguese children whose data were analysed used selected and adapted templates and that is independent of word size or word shape.

These preliminary results suggest that there are templates operating in the variability during phonological development. Not all children go through the same developmental path since variability and instability are inherent to development.

Discussion

Although children will eventually reach common ground in the phonology of their language, inter- and intra-variability are found in the way phonological development proceeds, as well as in transition instances. Portuguese and Brazilian children use the same strategy, namely reduplication, both within and between variety(ies), but word forms vary among children. Notice, however, that Brazilian children have templates during more age spans than Portuguese children due to the data collection in the corpora. Reduplication in child speech has specific forms and functions. It has been considered to be a trace of late babbling (Fee & Ingram, 1982) or 'bits of babble' (Lewis, 1936). Repeated syllables are interpreted as a means to facilitate the initial pronunciation (Fee & Ingram, 1982; Ferguson, 1983; Klein, 2005; Schwartz, Leonard, Wilcox, & Folgen, 1980). The production of two identical or partially identical syllables with phonological identity (Klein, 2005) has been reported previously for languages like English (e.g., [dɛdɛ] 'bye'/ [mima] 'grandmother' (Schwartz *et al.*, 1980). Repetitions of a core syllable have furthermore been associated with the production of a binary foot (Baia, 2008; Demuth, 1996; Santos 2007). In BP and EP, the production of reduplicated words with final prominence by young children is associated with the processing of iambic feet (Baia, 2008; Correia, 2010; Santos 2007). In this study, child-specific reduplication can be understood as the manifestation of a template/early articulatory routine. In sum, data from BP and EP phonological development show that children vary in the strategies they use to expand the lexicon as well as in the order of use (and disuse) of such strategies. Templates are instances of instability that characterize development and they are the result of self-organisation, that is spontaneous pattern formation.

Conclusion

Although we observed the use of reduplicated structures by Brazilian and Portuguese children, instability and variability were observed among children regarding the use and order of templates. Self-organisation instances of the phonological system differ, both within one child and between children. The system is self-organised due to its inherent capacity to form patterns from interaction. Therefore, variability appears as evidence of an open, dynamic and unstable phonological system.

Acknowledgments

This research was funded by *Fundação para a Ciência e a Tecnologia*, as part of the project of the Centro de Linguística da Universidade Nova de Lisboa – UID/LIN/03213/2013.

References

Baia, M. F. (2008). O formato prosódico inicial do Português Brasileiro: Uma questão metodológica? *Revel*, *6* (11).

Baia, M. F. (2010). A reduplicação fonológica na aquisição do português brasileiro: Uma comparação com outras línguas românicas. *Revel*, *8*, 31-46.

Baia, M. F. (2013). *Os templates no desenvolvimento fonológico: O caso do português brasileiro.* PhD Dissertation. Universidade de São Paulo.

Baia, M. F., & Correia, S. (2010). The initial prosodic template in Brazilian and European Portuguese: A methodological matter? In J. Costa, A. Castro, M. Lobo, & F. Pratas (eds.), *Language acquisition and development: Proceedings of GALA 2009* (pp. 13-47). Newcastle, UK: Cambridge Scholars Publishing.

Baia, M. F., & Correia, S. (2016). A auto-organização na emergência da fonologia: Templates na aquisição do português brasileiro e europeu. *Revel*, *14*(27).

Correia, S. (2010). *The acquisition of primary word stress in European Portuguese*. Unpublished PhD Thesis. University of Lisbon.

Davis, B., & Bedore, L. (2013). *An emergent approach to speech acquisition: Doing and knowing*. New York, NY: Psychology Press.

Demuth, K. (1996). The prosodic structure of early words. In J. Morgan & K. Demuth (eds.), *Signal to syntax*: *Bootstrapping from speech to grammar in early acquisition* (pp. 171-184). Mahwah, NJ: Lawrence Erlbaum.

Kelso, J. (1995). *Dynamical patterns: The self-organization of brain and behavior.* Cambridge, MA: MIT Press.

Fee, J., & Ingram, D. (1982). Reduplication as a strategy of phonological development. *Journal of Child Language*, *9*(1), 41-54.

Ferguson, C., & Farwell, C. (1975). Words and sounds in early language acquisition. *Language*, *51*(2), 419-39.

Ferguson, C. (1983). Reduplication in child phonology. *Journal of Child Language*, *10*(1), 239-43.

Keren-Portnoy, T., Vihman, M. M., DePaolis, R., Bidgood, A., & McGillion, M. (2011). The effect of production templates on word learning and word recognition. Presented at 12th IASCL, Montreal.

Klein, H. (2005). Reduplication revisited: Functions, constraints, repairs, and clinical implications. *American Journal of Speech-Language Pathology*, *14*, 71-83.

Lewis, M. M. (1936). *Infant speech: A study of the beginning of language.* New York: Harcourt, Brace & Co.

Macken, M. (1979). Developmental reorganization of phonology: A hierarchy of basic units of acquisition. *Lingua*, *49*, 11-49.

Santos, R. S. (2005). *A aquisição do ritmo em Português Brasileiro***.** Research project. University of Sao Paulo.

Santos, R. S. (2007). *A aquisição prosódica do português brasileiro de 1 a 3 anos: padrões de palavra e processos de sândi externo.* São Paulo: Universidade de São Paulo. Post-doctoral Dissertation.

Secco, G. (1994). *Criações lexicais em uma criança de 20 meses de idade.* MA Dissertation. Florianópolis: UFSC.

Schwartz, R., Leonard, L., Wilcox, M., & Folgen, K. (1980). Again and again: Reduplication in child phonology. *Journal of Child Language*, *7*, 75-88.

Szreder, M. (2012), The acquisition of consonant clusters in Polish: A case study. In M. Vihman & T. Keren-Portnoy (eds.), *The emergence of phonology: Whole-word approaches, cross-linguistic evidence* (pp. 343-362) Cambridge, UK: Cambridge University Press.

Thelen, E., & Smith, L. (1994). A *dynamic systems approach to the development of cognition and action.* Cambridge, MA: MIT Press.

Vihman, M., & McCune, L. (1994). When is a word a word? *Journal of Child Language*, *21*, 517-542.

Vihman, M., & Velleman, S. (2002). Whole-word phonology and templates: Trap, bootstrap, or some of each? *Language, Speech, and Hearing Services in* Schools, *33*, 9-23.

Vihman, M., & Croft, W. (2007). Phonological development toward a ‘radical’ templatic phonology. *Linguistics*, *45*(4), 683-725.

Waterson, N. (1971). Child phonology: A prosodic view. *Journal of Linguistics*, *7*, 179- 211.

To cite this article: Baia, M. F., & Correia, S. (2017). Self-organisation in phonological development: *templates* in Brazilian and European Portuguese. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 45-51). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

The acquisition of obstruent voicing contrasts of Eastern Armenian and Brazilian Portuguese as foreign languages

Maria de Fátima de Almeida Baia[1], Daniel Oliveira Peres[2]
mariadefatimabaia@uesb.edu.br, doliveiraperes@gmail.com

[1]Universidade Estadual do Sudoeste da Bahia (UESB), [2] Universidade de São Paulo (USP)

Abstract. This pilot study compares the acquisition of voicing contrasts by Armenian speakers from Yerevan learning Brazilian Portuguese (BP) with the acquisition of the same contrast by Brazilian speakers from Sao Paulo learning Eastern Armenian (EA). After analysing the data using perceptual-auditory and acoustic analysis, we observed that Armenian speakers produced correctly the BP contrast in initial and medial position but mistakenly kept the final contrast of EA in BP words ended with the grapheme *z* and did not apply the voicing rule between words. On the other hand, BP speakers did not produce the final word obstruent voicing contrast of EA and applied the BP voicing rule in EA sentences.

Keywords: Brazilian Portuguese, Eastern Armenian, L2 acquisition, obstruent, voicing

Introduction

This pilot study investigates the acquisition of voicing contrasts in Armenian and Brazilian Portuguese, as foreign languages. We compared the acquisition of voicing contrasts by Armenian speakers (Yerevan) learning Brazilian Portuguese (BP) with the acquisition of the same contrast by Brazilian speakers from Sao Paulo learning Eastern Armenian (EA).

We also investigated the *markedness differential hypothesis* (Eckman, 1987), which claims that relative degree of difficulty corresponds to relative degree of markedness: (i) the BP [s] and [z] voicing contrast in initial and medial word position; the voicing rule applied between words; (ii) the EA [s] and [z] voicing contrast in initial, medial and final word position; no voicing rule between words. According to the hypothesis, Armenian speakers would not report difficulty producing the BP voicing contrast as it occurs in the three positions of EA, whereas BP speakers would show more difficulty at performing the task, because there is no final voicing contrast in BP.

Therefore, this pilot study has the following main goals: 1) to compare the acquisition of voicing contrasts [s] and [z] by Armenian speakers learning Brazilian Portuguese (BP) with the acquisition of the same contrast by Brazilian speakers learning Eastern Armenian (EA); 2) to discuss briefly the notion of *markedness* in a *complex adaptive systems* (CAS) view (e.g., De Bot, 2008).

Armenian and Brazilian Portuguese: An overview

Armenian and Brazilian Portuguese are both Indo-European languages. Armenian is part of an independent branch of Indo-European languages. Due to the diaspora, Armenian has two linguistic varieties: Eastern Armenian and Western Armenian. Differences between these varieties are within phonological, morphological and lexical components.

Brazilian Portuguese is part of the Italic branch of Indo-European languages. Since the Portuguese colonization of Brazil, this variety has been undergoing changes in many linguistic respects, such as phonology, morphology and syntax.

Markedness and Eckman's hypothesis

Markedness is one of the core concepts in linguistic theory. It has been discussed in many studies with a focus on cross-linguistic analysis. Among linguists, Roman Jakobson (1972) and Troubetzkoy (1939) played a very important role introducing the markedness concept in their studies. A straightforward way to define *marked* vs. *unmarked* is to understand them as *usual* vs. *not expected* linguistic components. For instance, components such as long vowels, aspirated consonants are examples of marked linguistic units (e.g., Greenberg, 1966).

An important feature of markedness is predictability. Hume (2004) puts forward some properties commonly associated with unmarked sound patterns:

1. target of assimilation
2. result of neutralization processes
3. non-contrastive member of phoneme inventory
4. simple articulations
5. weak perceptual cues
6. more complex articulations in enhancement contexts, e.g. onset
7. strong perceptual cues in enhancement contexts
8. high predictability in a given context or system

Another important feature is frequency (Bybee, 2010). Based on this approach, high-frequency units have a stronger mental representation than low-frequency units. For this reason, they are easier to access and more available to either resist change or serve as its basis (Bybee, 2010:146).

Haspelmath (2006) divides markedness into four categories: markedness as *complexity*, as *difficulty*, as *abnormality*, and as a *multidimensional correlation*. The author also claims that the term markedness is superfluous and too general. Thus, it should be better expressed by less ambiguous terms. In other words, it should be replaced with primary explanations, or in the author's own words, '… simple everyday concepts should be expressed by simple everyday words' (Haspelmath, 2006:63).

Despite this criticism, this pilot study considers the *markedness differential hypothesis* (MDH) (Eckman, 1987:321). In L2 acquisition, markedness plays a role in the development of L2 grammatical features. In our view, although there is a debate on markedness as a linguistic term and concept, the MDH hypothesis is an important tool to predict what difficulties the L2 learners will face during the acquisition process. Briefly, MDH can be explained as: i) marked structures are more difficult to acquire than unmarked structures between two different languages; ii) L2 learners will have a better performance acquiring unmarked structures. In short, Eckman (1987) claims that the relative degree of difficulty corresponds to the relative degree of markedness.

In this study, this approach is suitable because there are different contexts where the voicing contrast occurs. Considering that both languages (BP and EA) have [s] and [z] in at least two different positions, we analyse if the position of the voicing contrast plays any role in its acquisition.

Methodology

Data collection

The audio data were collected from experiments performed by both groups of informants (Armenians and Brazilians): 8 participants, 4 Brazilians speakers from Sao Paulo learning EA and 4 EA speakers from Yerevan learning BP. Audios were recorded in 44kHz (.wav) using a ZOOM H2 digital recorder.

Experimental design

Two experiments were designed, one with BP words and sentences, and another with EA words and sentences. The participant's task was to read sentences and words in BP and AE. Four Armenians (AS1, AS2, AS3, AS4) acquiring BP (intermediate level) performed both experiments (EA and BP). Armenians performed the experiment in their own language in order to record native AE. These recordings were useful to compare with Brazilians' production. Four Brazilians (BS1, BS2, BS3, BS4) acquiring EA (intermediate level) performed the experiment with EA words and sentences. Both groups were aged between 20 and 30 years.

In BP, [s] and [z] occur in initial and medial position, whereas in EA they occur in initial, medial and final position. In a context s – z/___ [+voiced], the voicing process does not apply in AE, e.g., Ես ամերիկացի եմ [ie**s**.a.me.ɾi.ka.tsi.em] ‘I am American’. However, it applies for BP, e.g., *músicas alegres* [mu.zi.kɐ.**z**a.lɛ.gɾes] ‘happy songs’.

Table 1 summarizes the participant’s tasks:

Table 1. Participant’s tasks

Participants	Task	
Armenians	BP language: read 20 words, 10 sentences – context of s – z/___ [+voiced] between words – 7 words and 3 sentences as distraction stimuli.	EA language: read 24 words, 10 sentences – context of s – z/___ [+voiced] between words – 7 words and 3 sentences as distraction stimuli.
Brazilians		EA language: read 24 words, 10 sentences – context of s – z/___ [+voiced] between words – 7 words and 3 sentences as distraction stimuli.

The total number of words and sentences recorded was: i) Armenians reading EA: 96 words and 40 sentences; ii) Armenians reading BP: 80 words and 40 sentences; iii) Brazilians reading EA: 96 words and 40 sentences. In total, 272 words and 120 sentences were produced, recorded and analysed.

Data analysis and results

For the data analysis, we performed an auditory analysis and acoustic verification using Praat (Boersma & Weenink, 2013). As for findings for Brazilians’ reading EA, there was a considerable amount of accurate productions. On the other hand, similarly good performance was not observed when they read EA sentences in the context s – z/___ [+voiced].

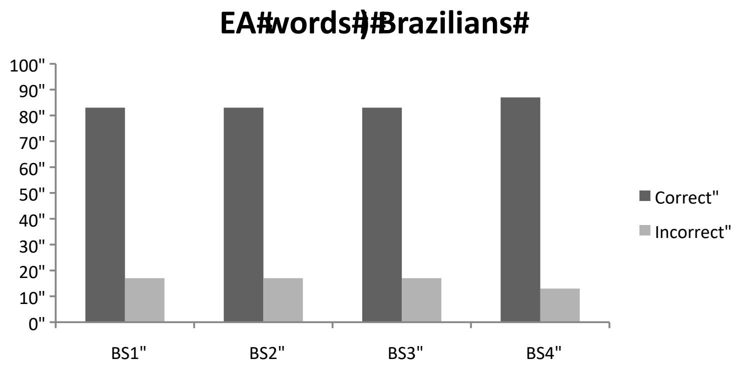


Figure 1. Results of Brazilians reading Armenian words

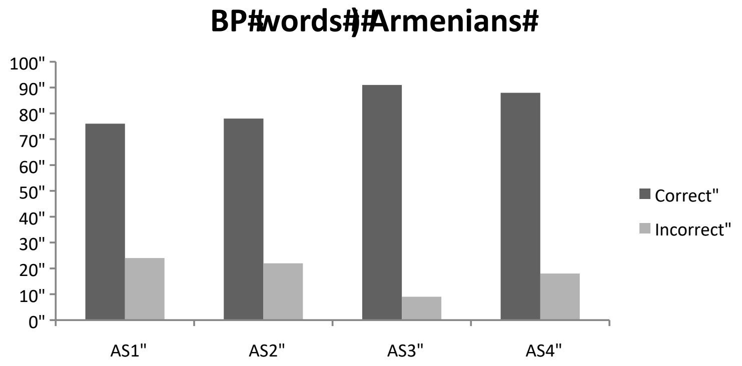


Figure 2. Results of Armenians reading Portuguese words

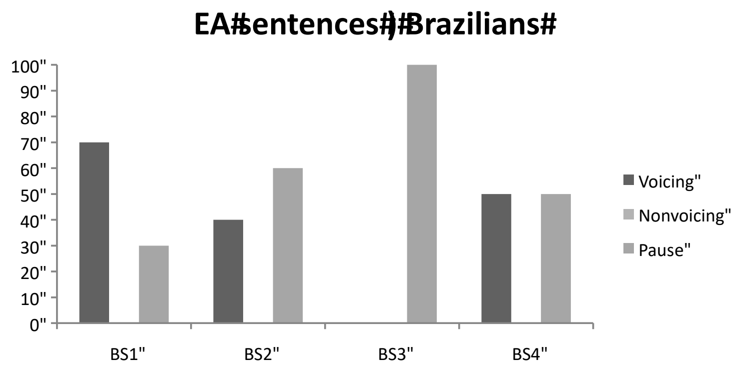


Figure 3. Results of Brazilians reading Armenian sentences

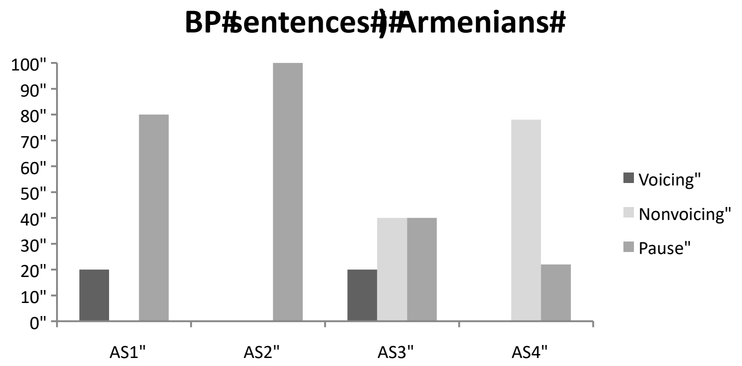


Figure 4. Results of Armenians reading Portuguese sentences

As we can see in Figures 1 and 2, Brazilians and Armenians performed the experiments fairly well reading words in BP and EA, although there were a few incorrect productions. Considering the Brazilian performance, they did not produce [z] in final position. On the other hand, Armenians mistakenly produced [z] while reading BP words with the final *z* grapheme in BP. It seems that it is difficult for Armenians to lose the word final contrast.

In Figure 3 (Brazilians reading EA sentences), the participants mistakenly produced the voicing rule between words as the voicing process does not apply to EA. The participants voiced [s] in final position, before vowels or voiced consonants. In addition, they produced a slight pause between the words. No [s] was produced by Brazilians in this context.

Armenians reading BP sentences (Figure 4) mostly produced [s] instead of [z] and pauses between the sentences in EA. The voicing rule was applied in a few cases (AS1 and AS3). Pause was the most common strategy used by EA native speakers speaking BP.

Some acoustic considerations

The data described previously underwent acoustic analysis. All the auditory results were corroborated by the acoustic analysis. However, we have some acoustic considerations regarding the final voicing contrast in Armenian.

The voicing in Armenian takes up a very small portion after the vowel, but the rest of the fricative does not present the voicing bar. One may argue whether the participant is partially voicing because Armenian has this distinction, or whether it is something inherent in the articulation of these sounds. As can be seen in Figure 5, there is a trace of glottal activity that runs out after the beginning of the fricative.

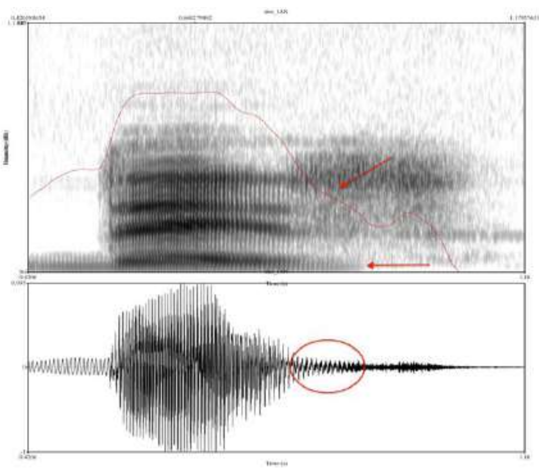


Figure 5. Armenian (AS1) production of the word Դեզ [dez] 'risk'

In the spectrogram (Figure 5), the top arrow indicates the loss of intensity that may be caused by the impedance applied by the vocal folds. The arrow below indicates the voicing bar, which will fade over time. In the waveform, the red circle corresponds to the parts indicated by the arrows in the spectrogram. In the acoustic signal, there are indications of periodicity, besides the clear energy loss, also indicated by the red line.

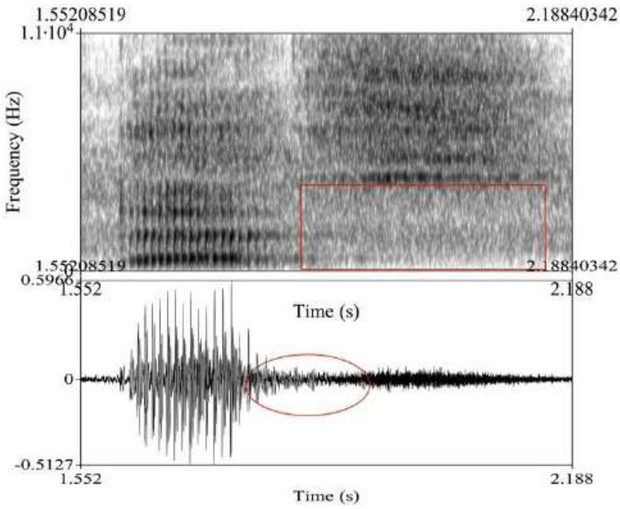


Figure 6. Brazilian (BS2) production of the word Դեզ [dez] 'risk'

In the spectrogram (Figure 6), the red rectangle shows that there is no voice bar. In the waveform, the red circle shows that there is no periodicity (Figure 6). This indicates that the vocal folds were not vibrating while the speaker was producing the fricative.

Conclusion

Markedness plays a role in the acquisition of a foreign language. This can be observed in the Brazilians' results of word final contrast in EA, since they did not produce [z] in final position. However, other aspects have to be taken into consideration in the analysis such as different prosodic contexts in addition to linguistic aspects. For instance, frequency, predictability and even historical adaptations (Briscoe, 1998) should be considered when markedness is under discussion. Moreover, orthography seemed to have an influence on the tasks performed by Armenians but not on those performed by Brazilians. This may be explained considering the differences between the Armenian and Latin alphabets. The Armenian alphabet shows more biunivocal relation between graphemes and phonemes than the Latin alphabet. In other words, although there are some exceptions, in general there is one letter for each phoneme. In contrast, the Latin alphabet has multiple relations between graphemes and phonemes. For instance, the Armenian alphabet has one corresponding letter for [s] <u>. The Latin alphabet, specifically in Brazilian Portuguese, [s] has multiple corresponding letters and digraphs <s, ss, c, ç, xc, xç, sc, sç>. We could argue that Armenians had difficulties reading BP because of the relation between grapheme/phoneme, but that does not explain why Brazilians did not read word final [z] despite the fact that there is only one <z> grapheme in Armenian. This is a reason why we intend to work on more data of this kind to present further results.

References

Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer* [Computer program]. Version 5.3.83, retrieved 11 May 2014 from http://www.praat.org.

Briscoe, T. (1998) Language as a complex adaptive system: coevolution of language and of the language acquisition device. In P. Coppen, P. Halteren & I. Teunissen. *Selected papers from the English CLIN meeting* (pp. 3-44). Amsterdam: Rodopi.

Bybee, J. (2010) Markedness: Iconicity, economy, and frequency. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology* (pp. 131-147). Oxford, UK: Oxford University Press.

De Bot, K. (2008) Second language development as a Dynamic System. *The Modern Language Journal*, 92, 166-178.

Eckman, F. R. (1987). Markedness and the contrastive analysis hypothesis. In **G.** Ioup & S. H. Weinberger (ed.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 55-69). Cambridge, UK: Newbury House Publishers.

Greenberg, J. H. (1966). *Universals of language* (2nd edition). Cambridge, MA: MIT Press.

Haspelmath, M. (2006). Against markedness (and what to replace it with). *Journal of Linguistics*, *42*, 25-70.

Hume, E. (2004). Desconstructing markedness. *Annual meeting of the Berkeley Linguistics Society Proceedings*, 182-198.

Jakobson, R. (1972) *Child language, aphasia and phonological universals.* Paris: Mouton.

Troubetzkoy, N. S. (1939) *Principles of phonology.* Translated by C.A.M. Baltaxe. Berkeley: University of California Press.

To cite this article: Baia, M. F. A., & Peres, D. O. (2017). The acquisition of obstruent voicing contrasts of Eastern Armenian and Brazilian Portuguese as foreign languages. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 52-57). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Connection between interactional context, parental discourse strategies and the child's morphosyntactic development in bilingual perspective. Evidence from a case study of a French-Russian bilingual child from 2;5 to 2;10

Oksana Bailleul
oksana.bailleul@gmail.com

University of Rouen Normandy

Abstract. The present longitudinal study examines the correlation between parental discourse strategies, language choice and morphosyntactic development of a child growing up with two languages from birth, French and Russian. The data were registered in two interactional contexts, joint book reading and free play from 2;5 to 2;10. The analysis supports the idea that the child's language choice in a weaker language (Russian) is highly correlated with context and parental discourse strategies: Russian utterances are more frequent in adult-directed speech during joint book reading. In this context, maternal monolingual strategies described by Lanza (2004) are more recurrent. French monolingual and mixed utterances are more frequent during free play, which has a strong correlation with the occurrences of bilingual strategies in adult speech. Parental discourse strategies show a positive correlation with the mean length of utterance (MLU) of the child's weaker language and mixed utterances. No correlation has been observed between these two variables and the MLU in French (stronger language) in dual-lingual conversations. This research supports the hypothesis of an impact of parental discourse-strategies on the child's language choice and the role of input frequency in bilingual simultaneous development at the early ages.

Keywords: bilingual child, MLU, parental discourse strategies

Introduction

In the past few decades, a growing interest in simultaneous bilingual development of young children has emerged. Recent research has shown that these bilinguals have different developmental paths at the beginning of grammatical development than monolinguals. An uneven development of one of the languages is often reflected in a greater amount of mixed utterances (Bernardini & Schlyter, 2004) and simple grammatical structures in a weaker language (Schlyter, 1995). Also, the length of the child's utterances in one language does not always relate to the length of his/her utterances in the other. To evaluate the child's language dominance, quantitative measurement of mean length of utterance (MLU) is widely used in combination with the upper bound (longest utterance) (Brown, 1973). This measurement is not without weaknesses, especially when two typologically distinct languages are compared, but it can show a clear idea of child's morphosyntactic skills in his two languages.

Family language policy studies have shown that parental language ideologies and beliefs have a significant impact on the bilingual children's linguistic outcomes (King, Fogle, & Logan-Terry, 2008; Lanza, 2004). In fact, parents with strong impact belief of their role in the harmonious bilingual development of their children have better chances of having children who speak two languages. This is due to the fact that these parents will seek linguistic opportunities in order to encourage children to speak both of their languages (De Houwer, 2009). Parental attitudes are also related to family rules regarding the use of mixed utterances by a child. Lanza (2004:261) argues that 'an examination of language mixing in child bilingualism requires a focus on parental discourse strategies towards this mixing, as this is an important aspect of the bilingual child's language socialization'. The use of monolingual and bilingual strategies by parents provides clues as to how the child perceives the context, as either monolingual or bilingual, and to which extend he can produce mixed utterances.

Parental monolingual strategies contribute to establishing bilingualism in the early years, especially in cases, where input in a non-societal language is restricted.

Thus, the purpose of this study is to explore the role of parental discourse strategies in the language choice and the MLU of both monolingual and bilingual utterances of a French-Russian bilingual child growing up with two languages from birth.

Data and methods

Participants

Despite language separation in the parental input, the child shows a clear dominance of French (societal and majority language at home). With her Russian-speaking mother, she engages constantly in dual-lingual conversations. The data are a part of a larger project which aims to examine the child's bilingual development from 2;1 to 4;0. For the present article, the child was recorded between 2;5 and 2;10 during joint book reading and free play. A total amount of 100 maternal and child turns were analysed within dual-lingual 'mother-child' conversations in each of these contexts and in three age categories: 2;6, 2;7-2;8; 2;8-2;9. The data were examined from three main perspectives which are the aim of the study: parental discourse strategies, child's language choice, quantitative measurement of MLU and upper bounds of French, Russian and mixed utterances.

Methods of measurement

The child's language choice was investigated within the framework of parental discourse strategies defined and discussed by Lanza (2004). Two bilingual strategies were analysed:

1. **Code switching**: the parent changes the language when the child produces a mixed utterance or a sentence in another language
2. **Move on strategy**: the parent doesn't use any of the below-listed monolingual strategies to correct the child's language choice when he produces a mixed utterance or a sentence in another language.

In a monolingual mode, we have analysed three strategies described by Lanza (2004):

1. **Minimal grasp strategy**: in response to a child's utterance in another language, the parent asks to translate or reformulate this sentence in a target language;
2. **Expressed guess strategy**: in response to a child's utterance in another language, the parent requests for clarification by reformulating it in a target language as a question;
3. **Repetition**: in response to a child's utterance in another language, the parent translates this sentence in a target language.

Alongside these, Ingold, Gendre, Rezzonico, Corlateanu, and Da Silva (2008) have identified *cognitive* discourse strategies, which aim to help the child achieve the task and get oriented within an ongoing activity and *linguistic* discourse strategies which concern linguistic form/function:

4. **Designation requests**: the parent asks the child to name an object in a target language;
5. **Designation proposals**: the parent names an object/action in a target language, when the child fails to name it in a previous turn;
6. **Clarification requests**: the parent asks for clarification of the child's intention;
7. **Self-repetition.**

Results

Parental discourse strategies

The analysis shows that despite the 'one person, one language' principle, the mother uses bilingual strategies in low proportions. The distribution of bilingual strategies varies according to the type of interactional context: code-switching is more recurrent in a book-reading context, but its rates decrease significantly for the overall period (see Table 1). In free play, these utterances are used only at 2;7-2;8 in very low proportions. In this context, the mother uses almost exclusively Move-On Strategy.

Table 1. Maternal bilingual strategies

	Joint book reading		Free play	
Age	**Code-switching (%)**	**Move-On Strategy (%)**	**Code-switching (%)**	**Move-On Strategy (%)**
2;6	2.26	3.01	0	8.24
2;7 – 2;8	0	3.03	0.39	7.0
2;8 – 2;9	0.27	6.79	0	7.88

This discrepancy is due to the context impact on the maternal verbal behaviour. In joint book reading, the adult pays more attention to the child's language use in both languages. She uses more code-switching in order to correct phonetics, vocabulary or morphology of the child's utterances in French. In free play, she is more focused on playing and, thus, uses Move-On Strategy more frequently.

With regard to monolingual strategies, their differential distribution depends to a great extent on context and the child's age.

Table 2. Maternal monolingual strategies in joint book reading and free play

Joint book reading (%)*								
Age	**Cognitive strategy**	**EGS***	**Designation Request**	**Designation Proposal**	**Self - repetition**	**Clarification Request**	**Repetition**	**Minimal Grasp Strategy**
2;6	23.31	15.04	14.29	9.78	12.78	8.27	5.16	3.76
2;7 - 2;8	29.70	18.18	13.94	9.00	13.94	9.70	1.82	0.61
2;8 - 2;9	48.64	12.23	10.05	6.50	9.78	5.16	2.99	0.27
Free play (%)*								
Age	**Cognitive strategy**	**EGS***	**Design. Request**	**Designation Proposal**	**Self-repetition**	**Clarification Request**	**Repetition**	**Minimal Grasp Strategy**
2;6	52.20	12.64	2.20	1.10	10.99	7.69	3.85	1.10
2;7 - 2;8	49.09	13.23	6.61	3.89	9.73	7.0	4.28	0.78
2;8 - 2;9	52.74	16.10	6.85	4.77	6.85	7.19	1.70	0

NB. The calculation is based on both monolingual and bilingual maternal strategies.
EGS***** = Expressed Guess Strategy.

In free play, the cognitive strategy is used in considerably higher proportions (see Table 2). This strategy helps the child organize, anticipate and become more conscious of her actions. During joint book reading, the adult uses this strategy in order to comment on actions, characteristics and mental states of the characters. Its occurrences show some major variations according to the child's age: at early ages, this strategy is substantially more recurrent in free play, whilst at 2;8 and 2;9, it is used in comparable proportions in both contexts. Clarification Requests, Repetition and Minimal Grasp Strategy are used in similar proportions in both contexts. A slight decrease of their occurrences is also observed from 2;6 to 2;9. Designation Requests, Proposals and Expressed Guess Strategy are more recurrent during joint book reading. The occurrences of the following strategies have also major variations according to the child's age: Clarification Request, Repetition, Self-Repetition and Minimal Grast Strategy in both contexts. As to Clarification and Designation Requests, their occurrences are less frequent at 2;8 and 2;9 during joint book reading (see table 2). This supports the idea that the mother is more focused on the language use during joint book reading, which helps to promote the child's weaker language. In free play, on the contrary, the adult is less focused on the child's language choice and her lexical/syntactical accuracy in both languages. The higher rates of the cognitive strategy in this context demonstrate that the adult's main scope is to help the child to get oriented within an ongoing activity. These observations show some major changes in the maternal verbal behaviour throughout the child's age span: at 2;8 and 2;9, the adult questions less the child's intention by clarifications, repetitions, suggestion as the child becomes more proficient in both languages. This idea is supported by the fact that the total number of monolingual Russian utterances has a permanent growth in the dual-lingual 'mother-child' conversations (see Table 3). Also, the child's utterances become more syntactically accurate at these ages (see Tables 4 and 5).

Child's language choice and sentence complexity

The child's linguistic competence in both languages is uneven: French MLU shows a stable increase, whilst Russian MLU has no significant growth from 2;1 to 2;10 (see Figure 1). A sudden increase in Russian MLU is observed at 3;3 which is positively correlated with linguistic input frequency. From this age and on, this measure shows a steady growth in Russian. Mixed utterances' MLU is significantly higher from 2;2 to 2;9, but their occurrences are extremely low from 3;3 and onwards. The breaking point of the language shift from dominant to harmonious use of both languages at about 3;0 is confirmed by the MLU measure: at this age, an increase of MLU in Russian is observed, even if MLU in French is slightly advanced. This discrepancy can be explained in terms of language specifics: Russian grammar has a synthetic-inflectional structure, while French is a moderately inflected language (see Figure 1).

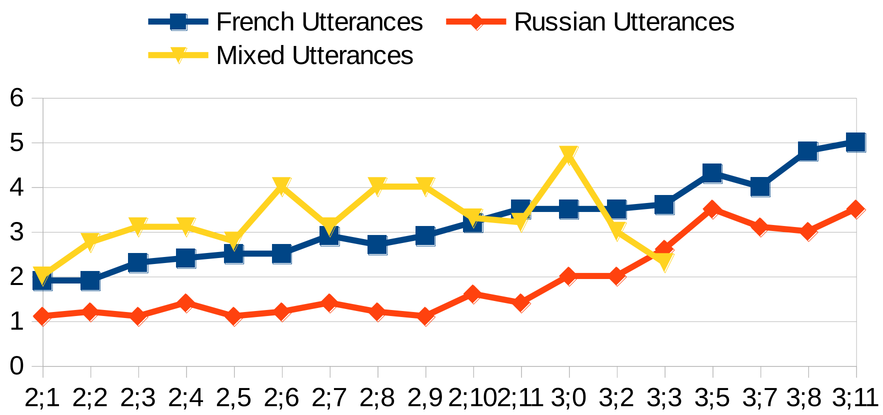


Figure 1. MLU of monolingual and bilingual utterances from 2;1 to 4;0

Our data support the idea that at the transitional period (shift from dominant bilingualism to harmonious use of both languages from 2;6 to 2;10), the interactional context plays an important role in the child's language choice. The total number of French utterances in the child's speech directed to her mother in dual-lingual conversations decreases in all contexts for the entire period (see Table 3).

Table 3. Child's language choice in joint book reading and free play

Joint book reading % (number of utterances)			
Age	**French Utterances**	**Russian Utterances**	**Mixed Utterances**
2;6	77.27 (51)	21.21 (14)	1.52 (1)
2;7-2;8	62.00 (31)	34.00 (17)	4.00 (2)
2;8-2;9	67.74 (63)	22.58 (21)	9.68 (9)
Free play % (number of utterances)			
Age	**French Utterances**	**Russian Utterances**	**Mixed Utterances**
2;6	82.58 (128)	11.61 (18)	5.81 (9)
2;7-2;8	63.10 (53)	27.38 (23)	9.52 (8)
2;8-2;9	66.25 (53)	28.75 (23)	5.00 (4)

The comparison of the total number of French, Russian and mixed utterances shows that during joint book reading, the child uses Russian with her Russian-speaking mother more frequently. At 2;6, these utterances constitute 21.2%, against 11,6% in free play. At further ages, this discrepancy is less observable, but the total number of Russian sentences remains still higher during joint book reading (see Table 3). A closer look at the data indicates that mixed utterances are more frequent in free play at all ages. This has a strong correlation with the total frequency of the bilingual Move-On Strategy used by the mother in this context and supports the hypothesis that it has a negative impact on the child's language choice. The data suggest that the child's MLU of monolingual Russian and mixed utterances depends on the context and the parental discourse strategy.

Table 4. MLU of French, Russian and mixed utterances (in words)

Joint book reading			
	2;6	2;7-2;8	2;8-2;9
French MLU	3	2.8	2.8
Russian MLU	1	1.1	1.1
Mixed Utterances MLU	3	3	4
Free play			
	2;6	2;7-2;8	2;8-2;9
French MLU	2.5	2.9	3
Russian MLU	1.2	1.3	1.6
Mixed Utterances MLU	4	3.5	4

Russian MLU is considerably higher in free play and has a steady growth in this context (see Table 4). In joint book reading, it shows no increase: this rate is over 1 word/utterance. This discrepancy is due to the higher rates of Designation Requests in the latter context, e.g. 14.2% against 2.2% in free play at 2;6 (see Table 2). The mother asks to name the book characters, objects or actions undertaken by these characters in Russian, which has a negative impact on the child's sentence complexity in a weaker language. As a result, the child switches either to French or uses Russian single-word utterances, which reflect her restricted competence in this language. Thus, Designation Requests have a positive effect on the child's language choice of Russian and, at the same time, a constraint effect on their MLU, which is due to her weaker proficiency in this language. In free play, lower rates of this strategy (2.2% and about 6% from 2;7 to 2;9) procure a less constrained effect on the child's MLU of Russian sentences (e.g., 1.6 word, against 1.1 word/utterance in joint book reading at 2;8-2;9). This tendency offers the child greater freedom to adjust her verbal behaviour, and in particular the sentence complexity in a weaker language to her communicative action and/or intention. Restricted maternal verbal guidance in this context, associated with higher rates of the cognitive strategy allows a child to use more initiative turns with higher syntactic complexity (see example 1).

Example 1

M3848: nautshi menya hodit'! govorit Masha. Pokazhi kak ty hodish?
'teach me how to walk, says Masha. Show me how do you walk?'
C4605: pomogay mama 'help (me) mother'
M3849: nautshite menia, ya ne umeyu hodit'! 'teach me, I don't know how to walk'
C4606: est tombée *(about the doll)* 'fell down'
M3850: ne paday mama *(to the doll)* 'don't fall down'
C4607: seytchas upadet! *(she) will fall down now)* '(she) will fall down now'

The MLU of French utterances, on the contrary, has no significant variance in these contexts and across the ages in dual-lingual 'mother-child' conversations (see Table 4). The discrepancy between Russian and French utterances' MLU in these contexts provides evidence that maternal use of either monolingual or bilingual strategies has no significant influence on the sentence complexity in the child's stronger language but procures a differential effect on Russian.

Moreover, higher rates of Move-On Strategy in free play have a positive impact, not only on the total number of mixed utterances in the child's speech, but also on their MLU. From 2;6 to 2;8, this strategy is used in higher proportions by the mother in free play (e.g., 7-8%, against 3% during joint book reading). The MLU of mixed utterances is also slightly higher (3.5 and 4 words/utterance, against 3 words in joint book reading). In the next age category (2;8-2;9), the adult uses this strategy in the same proportions in both contexts (7-8%) and the child's mixed utterances' morphosyntactic complexity remains the same (4 words/utterance). This tendancy shows that Move-On Strategy offers the child flexible verbal conditions in order to express complex ideas within dual-lingual conversations.

Another positive correlation has been observed between lower rates of maternal verbal guidance, higher rates of cognitive strategy and mixed utterances' MLU in the child's initiative turns (example 2).

Example 2
C3881: attention ɥu‿ɛl est? *(about the turtle)* où elle est, 'be careful, where is it?'
M3166: ishtshi tsherepahu. Gde ona? 'look for the turtle. Where is it?'
C3882: i s'habille Camille. I: nage Camille 'Camille is dressing up. Camille is swimming'
M3167: plyvi 'go swimming'
C3883: i cherche rybku Camille 'Camille is looking for a fish'

Table 5. Upper and lower bound (%)

Free play									
	French			Russian			Mixed		
	2;6	2;7-2;8	2;8-2;9	2;6	2;7-2;8	2;8-2;9	2;6	2;7-2;8	2;8-2;9
Upper bound	6 words (7.0)	7 words (2.0)	7 words (2.0)	2 words (22.2)	3 words (6.25)	2 words (26.0)	6 words (11.1)	5 words (28.5)	3 words (50.0)
Lower bound	1 word (28.3)	1 word (12.2)	1 word (12.2)	1 word (77.7)	1 word (68.7)	1 word (73.9)	2 words (14.2)	2 words (50.0)	2 words (50.0)
Joint book reading									
	French			Russian			Mixed		
Upper bound	6 words (13.3)	6 words (16.3)	7 words (1.5)	1 word (100)	3 words (7.1)	3 words (4.3)	3 words (100)	3 words (50.0)	6 words (16.5)
Lower bound	1 word (16.6)	1 words (12.9)	1 word (23.8)	*idem*	1 word (78.5)	1 word (91.1)	*idem*	2 words (50.0)	2 words (12.5)

The upper bound analysis confirms that the MLU is significantly higher in free play (both languages). The Lower Bound occurrences are less frequent in this context in comparison with joint book reading. Mixed-utterances Upper Bound is also substantially higher in free play. In this activity, a greater variety of multi-word mixed utterances is observed. Upper Bound of French utterances is similar in both contexts, but the occurrences of these multi-words utterances are substantially higher during joint book reading, especially from 2;6 to 2;8 (see Table 5).

This discrepancy confirms the role of input frequency of both languages. During the recordings, both parents read the same book to the child in two languages. Thus, the child receives a similar proportions of input related to the story and the girl remains in bilingual mode, in terms of Grosjean (2015). At the same time, the father engages less in free play. The input related to this context is thus more diverse in Russian.

Discussion

The results of the present study allow us to validate the existence of a positive correlation between the interaction context, parental discourse strategies, the child's language choice and the MLU in dual-lingual 'mother-child' conversations. Lanza (2004) has argued that parental discourse strategies are of primary importance for the harmonious development of young bilinguals. Our results confirm that these strategies play an important role, especially in the context where one of the languages is not spoken in the community and its input is limited.

The analysis has shown an uneven development of two languages with a strong dominance of French (the societal language). The child's choice of Russian positively correlated with the mother's use of monolingual strategies (Designation Requests, Proposals, Expressed Guess Strategy, Minimal Grasp Strategy) during joint book reading. French and mixed utterances are more frequent in free play at all ages, which has a strong correlation with maternal use of Move-On Strategy in this context. These results support the idea that this strategy has a negative impact on the child's language choice of the weaker language. Moreover, the child's sentence complexity has major variations influenced by the interaction context and parental discourse strategies. Within dual-lingual conversations, French MLU has no significant variations according to these interactional parameters and across the ages. This is due to the child's stronger proficiency in this language. However, these two variables (context and discourse strategy) have profound impact on Russian and mixed-utterances: their MLU is considerably higher in free play, which has a steady growth in this context at all ages. In free play, lower rates of maternal verbal guidance (Clarification and Designation Requests, Repetition, Minimal Grasp Strategy, Expressed Guess Strategy), associated with higher proportions of cognitive strategy procure the child a possibility to use more initiative turns with a higher syntactic complexity. Relatively low proportions of code-switching in maternal speech didn't allow us to discuss Lanza's hypothesis about the negative impact of this strategy on the child's language choice.

Conclusion

This study extends our understanding of the role of interactional factors in the harmonious bilingual development of young children growing up with two languages from birth. The present study demonstrated that parental and child interactional behaviours depend on the context: in joint book reading, the adult is more focused on the child's language use, which allows to promote the weaker language. The child's language choice of the weaker language (Russian) is highly correlated with the maternal monolingual discourse strategies. Also, our results provide evidence that parental strategies produce strong effect on the MLU of the child's utterances, especially in the weaker language and in mixed utterances.

References

Bernardini, P., & Schlyter, S. (2004). Growing syntactic structure and code-mixing in the weaker language: The Ivy Hypothesis. *Bilingualism: Language and Cognition, 7*(1), 49-69.

Brown, R. (1973). *A first language: The early stages*. London: George Allen & Unwin.

De Houwer, A. (2009). *Bilingual first language acquisition*. Clevedon: Multilingual Matters.

Döpke, S. (1998). Competing language structures: The acquisition of verb placement by bilingual German-English children. *Journal of Child Language*, 25(3), 555-584.

Grosjean, F. (2015). *Parler plusieurs langues. Le monde des bilingues*. Paris: Albin Michel.

Ingold, J., Gendre, S., Rezzonico, S., Corlateanu, C., & Da Silva, C. (2008). Diversité des étayages des mères d'enfants tout-venant et dysphasiques dans deux situations d'interaction. *Revue Tranel (Travaux neuchâtelois de linguistique)*, *49*, 69-82.

King, K.A., Fogle, L., & Logan-Terry, A. (2008). Family Language Policy. *Language and Linguistics Compass, 2*(5), 907-922.

Lanza, E. (2004). *Language mixing in infant bilingualism: A sociolinguistic perspective.* Oxford, UK: Oxford University Press.

Schlyter, S. (1995). Formes verbales du passé dans des interactions en langue forte et en langue faible. *Acquisition et Interactions en Langue Etrangère*, *6*, 129-52.

To cite this article: Bailleul, O. (2017). Connection between interactional context, parental discourse strategies and the child's morphosyntactic development in bilingual perspective. Evidence from a case-study of a French-Russian bilingual child from 2;5-2;10. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 58-64). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

More evidence for selective attention to features: Perception of the Turkish high back unrounded vowel

Anna Balas
abalas@wa.amu.edu.pl

Adam Mickiewicz Univeristy in Poznań

Abstract. The evidence for selective attention to features comes from Bohn & Best (2012) and Pajak & Levy (2014) and refers to the features known from the L1 and re-employed in the non-native language. Balas (2017) shows that L1 Polish listeners can more faithfully perceive front rounded vowels as front rather than back vowels if they know an L2 which uses front rounded vowels. This study examines whether learners of languages with the distinctions between front unrounded vs. front rounded vowels can also perceive two Turkish vowels more faithfully: a high back unrounded vowel /ɯ/ and a mid-front centralized rounded vowel /œ/. A perceptual assimilation task was administered. Polish learners of English, French or Dutch were asked to match the Turkish vowels they heard with the Polish vowel labels, and further to indicate the goodness of fit. The results revealed that Polish learners of English were not able to assimilate the Turkish high back unrounded vowel to any of the Polish vowels, whereas Polish learners of French and Dutch could perceive the height of this vowel and more consistently assimilate it to Polish front or back vowels. This is interpreted as the ability to disentangle backness from rounding thanks to abstracting the feature [+rounded] which has a distinctive function in French and Dutch.

Keywords: non-native vowel perception, Turkish, high back unrounded vowel, mid-front centralized rounded vowel

Introduction

So far, L2 speech perception has been viewed in terms of assimilations of these L2 sounds which are similar to L1 sounds and in terms of new category formation for definitely different sounds (Best, 1995; Best & Tyler, 2007; Flege, 1995). Bohn and Best (2012), however, present findings about increased sensitivity to non-native contrasts that differ in features contrastively used elsewhere in the L1. Pajak and Levy (2014), therefore, postulate an important role for selective attention to features which, in addition to categories, might be used in non-native speech perception even if the context for using these features is different.

In Bohn and Best's (2012) study, Danish and German listeners, whose native approximant system has only three categories in a syllable-initial position (/r, l, j/), discriminated American English /w - j/ comparably to French listeners and exceedingly better than native American English listeners. This means that structural phonological and allophonic differences related to approximants between the three languages do not account for the results. Bohn and Best (2012) notice however that what French, Danish and German share, in opposition to English and Japanese, is a systematic distinction between front rounded and unrounded vowels. The distinctive, contrastive lip rounding feature for vowels is used in French, Danish and German, but it is absent both phonologically and phonetically from English and Japanese vocalic systems. Semivowels /j/ and /w/ are actually short, non-syllabic versions of the unrounded /i/ and rounded /u/. Bohn and Best (2012) propose that high sensitivity to lip rounding distinctions in native vowels allows Danish and German listeners to discriminate a corresponding non-native approximant contrast, as this contrast is distinctively differentiated by lip rounding.

Non-native speech perception can be interpreted as a process involving 'learned patterns of selection and integration of those acoustic properties of speech stimuli that are phonologically relevant in the native language' (Strange & Schafer, 2007). Pajak and Levy (2014) postulate an important role of selective attention to features in L2 speech perception. Selective attention to features may be used to account for problems in discriminating the contrasts that differ along unattended dimensions, because L1 perceptual weighing of the relevant acoustic and articulatory cues will influence non-native speech perception. On the other hand, non-native speech perception might be fostered if the cues that are relevant for the L1 can be used for discrimination in the non-native language, even if the enhancement should apply in different acoustic and phonetic contexts in the foreign language. The idea that cues are reused dates back to Clements (2003). He argues that languages reemploy a limited set of phonetic features for signaling multiple contrasts. Consequently, it seems that it is more economical for the listener to enhance sensitivity to features which can be reused in numerous contexts rather than to enhance sensitivity to specific L1 categories. In Pajak and Levy's (2014) study, listeners with L1s which have vowel length distinctions performed better on consonant length distinctions. Pajak and Levy (2014) conclude that non-native speech perception is mediated not only by L1 phonetic categories, but also by more general phonological principles. The issue now is to determine what exactly these phonological principles are and how they influence speech perception.

Balas (2017) tested whether there is a language-specific effect of L2 acquisition on foreign vowel perception and whether the acquisition of the L2 with specific vowel features does or does not facilitate the perception of different foreign language vowels with the same feature (cf. Bohn & Best, 2012, Escudero & Williams, 2012). Specifically, it was examined how the feature [+ rounded] in various familiar configurations (universal back rounded vowels present in Polish, English, though in a centralized form, French and Dutch) and unfamiliar configurations (front rounded vowels, which are present in Dutch and French, but not in Polish or in English) influences perception of foreign vowels. Dutch front rounded vowels were identified predominantly as front vowels by learners of French and Dutch and as back vowels by learners of English. The results suggest that the experience with second language front rounded vowels is enough to trigger disentangling rounding from backness and perceiving the frontness of the vowels, despite their rounding. Balas (2017) study however did not reveal whether the Dutch and French learners succeeded in perceiving the Dutch front rounded vowels as front vowels, because:

1. they were simply familiar with these vowels from their L2s: they managed to perceive the more marked front rounded vowel and counteract the universal hierarchy, in which rounding implies backness (see Crothers, 1978)
2. of selective attention to features at play here. Thanks to their experience with front unrounded and rounded vowels in their L2s learners could learn to actively use the feature [+rounded].

In order to verify which hypothesis is correct, the present study examines the perception of unfamiliar Turkish vowels: a back unrounded vowel /ɯ/ and a mid-front centralized rounded vowel /œ/. If the first hypothesis is true, the perception of the Turkish vowels, especially of the /ɯ/, will be similar by Polish learners of English, French and Dutch. If the second hypothesis is true, French and Dutch learners will perceive the Turkish vowels more consistently than English learners.

Method

Subjects

The subjects were native speakers of Polish who learn English, French or Dutch as their L2. All subjects were 2nd year university majors in a respective language. Second language learners rather than native speakers of English, French or Dutch were chosen, because the subjects for the three groups could categorize foreign language sounds to their L1 categories. This way, a common L1 served as *tertium comparationis* for assimilations potentially mediated by the second language. There were 17 listeners with L2 English, 10 listeners with L2 French, and 25 listeners with L2 Dutch.

Stimuli

The recordings took place in a studio with an anechoic chamber equipped with digital audio recording gear. The stimuli were recorded by a male native speaker of Turkish. The speaker read sentences with eight Turkish vowels /i y ɯ u e œ o a/ in a carrier sentence: *SVs içinde V bulunur*. Only words with /ɯ/ and /œ/ were used for the present experiment. Additionally, eight Dutch *sVs* words, recorded by a native speaker of Northern Standard Dutch, were used in the experiment as fillers.

Procedure

The experiment consisted of an identification test with category goodness ranking of Turkish vowels (and additionally Dutch vowels) in terms of Polish vowels. The subjects performed a keyword identification of the two Turkish vowels /ɯ œ/ and Dutch fillers in terms of six Polish vowel or vowel plus glide categories /i, ɨ, ɛ, a, o, u/. The labels in the experiment were orthographical: *i, y, e, a, o, u, ej,aj, ij, uj, eł, ał, ił, uł* since Polish vowel orthography is transparent. The tested items were ten vowel categories, three tokens per category, and five repetitions per token, which yields 150 trials per subject. The stimuli were presented randomly, and they were preceded by warm-up items. The subjects clicked on a Polish keyword corresponding to the vowel that they heard in the item and then rated the similarity of the vowel in the auditory stimulus to the vowel in the chosen keyword (1 being barely similar and 7 being a very good fit).

Results

Tables 1 and 2 present the categorization results for the two Turkish vowels /ɯ/ and /œ/ by Polish L2 learners of English, French and Dutch in terms of Polish vowel categories.

Table 1. Mean per cent categorization and goodness rating in parentheses of the Turkish high back unrounded vowel /ɯ/ by Polish L2 learners of English, French and Dutch in terms of Polish vowel categories and the response keywords (only the results above 3% are presented). The most frequently chosen identification response per target is boldfaced, whereas the next most frequent choice is italicized. The goodness ratings were based on a scale that ranged for 1 (barely similar) to 7 (identical)

L2	ɨ	e	a	o	u
English	**27.8%** (3.4)	7.45% (2.1)	20% (2.9)	20.4% (4.6)	*24.3%* (3.4)
French	9.3% (2.7)	*16%* (3.5)	15.3% (4.2)	**56.7%** (5.7)	
Dutch	**53%** (3.1)	*27.2%* (4.3)	11.47% (3.4)	2.13% (3.8)	5.07% (4.6)

Table 2. Mean per cent categorization and goodness rating in parentheses of the Turkish mid-front centralized rounded vowel /œ/ by Polish L2 learners of English, French and Dutch in terms of Polish vowel categories and the response keywords (only the results above 3% are presented). The most frequently chosen identification response per target is boldfaced, whereas the next most frequent choice is italicized. The goodness ratings were based on a scale that ranged for 1 (barely similar) to 7 (identical)

L2	ɨ	e	a	o	u
English	*21.6%* (3.2)	**32.2%** (3.5)	12.2% (3)	13.7% (4.3)	16.9% (3.4)
French	4.7% (1.7)	*38.7%* (4.5)	13.33% (4.2)	**40.7%** (5.6)	
Dutch	*43.7%* (3.2)	**44.8%** (4.8)	3.7% (2.7)		5.3% (4.8)

A repeated-measures ANOVA was run on all Dutch, English and French groups' arcsine-transformed correct identification percentages with Vowel as within-subject factor (10 levels, calculations performed for both the Turkish vowels reported here and the Dutch vowels) and L2 as between-subjects factor (3 levels). There was a main effect of Vowel ($F= 31.855$; $p< 2e-16$), indicating that some vowels were more difficult to perceive than others. There was a main effect of L2 ($F= 20.95$; $p=2.67e-07$), which means that there are significant differences between vowel identification between the three L2 groups. There was a significant Vowel x L2 interaction ($F=3.597$; $p<1.17e-06$) (more than for the Vowel, and slightly more than for the L2), which suggests that the three groups differ in their identification responses for some Dutch vowels. Table 3 shows the results of t-tests for /ɯ/. The differences for /œ/ were not statistically significant.

Table 3. T-test results showing whether the two chosen groups of L2 learners assimilated the /ɯ/ vowel differently. The results were contrasted with the most frequent answer given by the learners of Dutch. Statistically significant results (p-value<0.05) are given in bold

t-test values	Dutch - English	Dutch - French	English - French
t	**2.187**	**4.0934**	1.6765
p-value	0.03501	0.00032	0.1063

Discussion

In Balas (2017), Polish listeners with L2 English perceive Dutch high front rounded vowels predominantly as Polish /u/. In the present study, Turkish high back unrounded /ɯ/ and mid-central rounded /œ/ were examples of uncategorized foreign vowels (adopting a 50% threshold after Faris, Best, &Tyler, 2016). Polish listeners with no experience with front rounded vowels seemed not to have known how to categorize a back unrounded vowel in the presence of conflicting cues, i.e. backness and the lack of rounding or mid-front centralized rounded vowel, where rounding coincided with unfamiliar frontness and advancement of the tongue. Interpreting Balas (2017) and the present results together, we can speculate that for Polish listeners unfamiliar with front rounded vowels, a high rounded vowel implies a back vowel /u/, but a high back vowel does not entail a rounded /u/. Consequently, it seems that disentangling rounding from backness is difficult for Polish listeners in such a way that a feature [+ rounded] is associated with backness, but backness is not associated with rounding. Coupling of the features is therefore asymmetrical.

Learners of L2 French and Dutch, however, seemed to be able to disentangle rounding from backness in front rounded vowels because of their knowledge of French or Dutch (Balas, 2017). In the present study, the learners of French assimilated the /ɯ/ vowel to the back /ɔ/ (56.67%, goodness 5.73), whereas the learners of Dutch assimilated it to high and high-mid front vowels: /ɨ/ (53.07%, goodness 3.08) and /e/ (27.2%, goodness 4.34). The vowel was categorized (cf. Faris *et al*., 2016) by subjects familiar with front rounded vowels from their L2, unlike in the case of learners of English. The /œ/ vowel was not categorized, but at least in the case in the case of French and Dutch learners preferences where observed: learners of French decided to assimilate the vowel to Polish mid-vowels (both Polish and French have only basic frontness-backness distinctions), whereas learners of Dutch to Polish front high-mid and mid vowels. In the absence of a Polish counterpart, they decided to choose categories with one of the other features of the vowel: either height or centralized frontness. It seems that learners of French and Dutch could cope with categorizing the /ɯ/ vowel, because they can truly disentangle rounding from backness due to their experience with the feature [+rounded] in a contrastive function in French or Dutch. The present results can therefore be claimed to provide preliminary support for selective attention to features at play in foreign/L2 vowel perception in the case of contrastive features, be it in the L1 or the L2.

With regard to the hypotheses, the first one was not supported by the results. The Turkish /ɯ/ vowel, a new vowel for all the subjects, was not perceived in a similar fashion by the subjects from the three groups. We can observe partial support for the second hypothesis, as English learners were not able to categorize the Turkish vowels, whereas the learners of Dutch and French were able to associate them more consistently with Polish vowels. One could potentially expect that learners of Dutch and French, as those listeners who are experienced with front rounded vowels, would assimilate the foreign vowels in a similar vein. Yet, as the results show, it was not the case. Turkish /ɯ/ was categorized as /ɔ/ by the learners of French and as /ɨ/ by the learners of Dutch. Turkish /œ/ was categorized as Polish /e/ and /ɔ/ by learners of French and as or /e/ or /ɨ/ by learners of Dutch (non-significant difference between /e/ categorizations). It needs to be pointed out, however, that what was at play was not pure selective attention to abstract features, which could yield identical results for both groups of learners. What determines perception is the interplay of features and phonetic categories with their language-specific acoustic correlates (cf. Second Language Linguistic Perception model by Escudero & Boersma (2004) and Escudero (2005), who assume that listeners' perceptions of native and non-native sounds match the acoustic properties of relevant sounds in their native dialects).

Conclusion

This paper examined Turkish vowel perception by Polish learners of English, French or Dutch. The results showed that the sounds were more consistently assimilated by learners of French and Dutch who, thanks to the L2s they knew, were experienced in using the feature [+rounded] in a contrastive function. It is suggested that the experience with front rounded and unrounded vowels allowed them to disentangle rounding from backness and reliably perceive other features of the sound such as height or backness. The results also provided evidence of L2-specific effects in non-native speech perception. Learners of English were not able to categorize the Turkish vowels, whereas French and Dutch learners chose to opt for mid-high vowels – their answers were not scattered throughout the inventory. Finally, the results provided evidence of the interplay of selective attention to features, which have a contrastive/distinctive function in L2 and phonetic categories with their language-specific acoustic correlates. Further research should verify the findings presented here on the basis of larger samples of subjects, other phonetic features, and other language combinations.

Acknowledgements

I would like to thank Katarzyna Dziubalska-Kołaczyk, Magdalena Wrembel, Dafydd Gibbon, Jarosław Weckwerth and Brechtje Post for their advice either on the design of the study or comments after the presentation. I am also grateful to the native speakers who recorded the stimuli and participants in the experiments.

The project has been financed by grant no. UMO-2015/17/B/HS2/01246 'Perceptual reorganization of speech: the interplay of categories and features' from the National Science Centre, Poland.

References

Balas, A. (2017). The influence of second language vowels on foreign language vowel perception. *Linguistic Society of America Proceedings, 2*(44), 1-5.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ed.), *Speech Perception and linguistic experience: Issues in cross-language research* (pp. 171-204). Baltimore, MD: York Press.

Best, C. T., & Tyler, M.D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. Munro & O-S. Bohn (eds.), *Second language speech learning* (pp. 13-34). Amsterdam: John Benjamins Publishing Company.

Bohn, O-S. & Best, C.T. (2012). Native language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *Journal of Phonetics, 40*, 109-128.

Clements, G. N. (2003). Feature economy in sound systems. *Phonology, 20*, 287-333.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Utrecht: LOT.

Escudero, P., & Boersma P. (2004). Bridging the gap between L2 speech perception and phonological theory. *Studies in Second Language Acquisition, 26*, 551-585.

Escudero P., & Williams, D. (2012). Native dialect influences second-language vowel perception: Peruvian vs. Iberian Spanish learners of Dutch. *Journal of the Acoustical Society of America, 131*(5), EL406-EL412.

Faris, M. M., Best C. T., & Tyler M.D. 2016. An examination of the different ways that non-native phones may be perceptually assimilated as uncategorized. *The Journal of the Acoustical Society of America*, *139*(1), EL1-EL5.

Flege, J. E. (1995). Second language speech learning: Theory, problems, findings. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross- language research* (pp. 233-277). Baltimore, MD: York Press.

Pajak, B., & Levy, R. (2014). The role of abstraction in non-native speech perception. *Journal of Phonetics, 46*, 147-160. DOI: 10.1016/j.wocn.2014.07.001

Strange, W., & Schafer, V. L. (2007). Speech perception in second language learners: The re-education of selective perception. In J. G. Hansen Edwards & M. L. Zampini (eds.), *Phonology and second language acquisition* (pp. 153-191). Amsterdam: John Benjamins Publishing Company.

To cite this article: Balas, A. (2017). More evidence for selective to features: Perception of the Turkish high back unrounded vowel. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 65-70). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

The Sonority Syllable Model reconsidered: Two challenges and their solution

Hans Basbøll
hba@sdu.dk

University of Southern Denmark

Abstract. According to Ohala (1992) and Ohala & Kawasaki-Fukomori (1997), sonority or strength hierarchies are, in general, circular. I have attempted (e.g., Basbøll, 2005:173-201) to avoid this circularity by deriving the *Sonority Syllable Model* from the universal logic of segment types, by the implicational chain [+vocoid] > [+sonorant] > [+voiced] > [–spread glottis]: all vocoids are, necessarily, sonorant (by the definition [+vocoid] $=_{\mathrm{DEF}}$ [+sonorant, –lateral, –stop]), etc. The weakest point in this chain is [voiced], due to the occurrence of voiced aspirated plosives. A further challenge is that sonorant laterals and nasal contoids must have the same position in the chain since the feature distinguishing them, viz. [stop] (or [continuant]), cannot be placed into this implicational chain: segments that are blocked in the oral cavity can be either sonorants, viz. nasal contoids, or obstruents, viz. plosives. I have given (Basbøll, 2005:117-127) several proposals to depict major classes in a 2-dimensional space, by combining the sonority-hierarchical features and [stop]. I develop this approach further here, by combining the sonority-hierarchical features [vocoid], [sonorant] and [spread glottis] with the feature [stop] (which is part of the definition of [vocoid]). This gives an attractive strength or sonority hierarchy that is *not circular*. Particularly for obstruents, this hierarchy gives some interesting predictions: *initially*, /s/ plus unaspirated plosive is predicted; *finally*, unaspirated plosive plus /s/ is predicted; and if a plosive is more marginal than /s/ (e.g., *ps-, -sp*), it is predicted that the plosive is [+spread glottis]. Phonetically, anything else would be very surprising, and the glottis position can be observed, so the proposal can be tested.

Keywords: sonority, strength hierarchy, major classes, syllables, features

Introduction (1)

According to John Ohala (1992, 2008; Ohala & Kawasaki-Fukomori, 1997), sonority or strength hierarchies are, *in general*, circular. The basic argument is: *such hierarchies are based upon (generalizations from) observed phonotactic patterns and can therefore not explain phonotactic patterns*. I have attempted (since Basbøll, 1973, cf. 2005:173-201) to avoid this circularity by deriving the Sonority Syllable Model from the universal logic of segment types, by the implicational chain:

(1) [+vocoid] > [+sonorant] > [+voiced] > [–spread glottis]

All vocoids are, *necessarily*, sonorant, due to the definition [+vocoid] $=_{\mathrm{DEF}}$ [+sonorant, –lateral, –stop], and so on for the further ‘sonority-hierarchical features’ as I call them, viz. those mentioned in formula (1). But there are challenges to this model: The weakest point in this chain is [voiced], due to the occurrence of voiced aspirated plosives. And a further challenge is that sonorant laterals and nasal contoids must have the same position in the chain: *the feature distinguishing them, viz. [stop] (or [continuant]), cannot be hierarchical since segments which are blocked in the oral cavity can be either sonorants, viz. nasal contoids, or obstruents, viz. plosives.* And there is much evidence that sonorant laterals have a higher sonority than nasal contoids.

Constructing the model (2)

The vocoid as the prototypical peak of a syllable

All languages have vocoids as syllabic peaks, only some have non-vocoids = contoids as peaks (such as Czech *vlk* ‘wolf’, and Slovenian *krk* ‘neck’); all languages have contoids as non-peaks, only some have vocoids (glides, semivowels); thus, *vocoids are prototypical peaks*; since the peak-function is central in the notion of the syllable, the point of departure here is the *vocoid*. But it is important, to avoid circularity, that one should not conceive of vowels here in any functional sense (which implies peak-ness).

Ladefoged (1971:91) aptly says about his feature *Consonantal*:

> 'This feature has a different status from all other features in that it can be *defined* only in terms of the intersection of classes already defined by other features. Thus nonconsonantal sounds are nonlateral and sonorant [they are also oral/HB]. They correspond largely to what Pike (1943) called vocoids, which he defined as central resonant orals'.

In my view, such *cover features* (in Ladefoged's term) *are preferable* to independently defined features, other things being equal, due to Occam's razor principle; this is particularly true for major class features. The central *definition of the vocoid* is given in formula (2):

(2) [+vocoid] $=_{\mathrm{DEF}}$ [+sonorant, –stop, –lateral]

The features employed here are all strictly binary. Vocoids constitute a phonetically *homogeneous class*, their opposite member (contoids, according to Pike's terminology) does not, since they include plosives and fricatives as well as sonorant laterals, for example.

Sonorants are defined acoustically, following Ladefoged (1971:58) as having: 'a comparatively large amount of acoustic energy within a clearly defined formant structure' (cf. p. 93: 'greater acoustic energy in the formants'); they are, like their complementary class (obstruents), incidentally, phonetically homogeneous.

Universal logic of segment types: sonority-hierarchical features

There are some important consequences of the definition of the vocoid (formula 2):

The point of departure is the prototypical syllabic peak, which is a vocoid (a phonetic, as opposed to 'functional', vowel; as noted above, in the latter sense, this criterion would be *circular*): All vocoids are, necessarily, sonorant – this follows from the definition. But some sonorants are not vocoids, viz. prototypical (sonorant) laterals (i.e. [+sonorant, +lateral]) and nasal contoids (i.e. [+sonorant, +stop]).

Ergo: [+vocoid] IMPLIES [+sonorant] (and not the other way round)

Furthermore, all sonorants are, *necessarily*, voiced: this follows from the definition used here (Ladefoged, 1971:58, 93) combined with the phonetic (articulatory and acoustic) fact that in order to obtain sufficient acoustic energy in the formants, the vocal chords must vibrate; this is especially critical in the case of F1, because the intensity of higher formants decreases with frequency. On the other hand, there are non-sonorant sounds (called obstruents) that are voiced.

Ergo: [sonorant] IMPLIES [voiced] (and not the other way round)

The implications of this argument so far can be depicted by means of a set of Euler's circles as in Figure 1 (from Basbøll, 2005:182).

Furthermore, all [+voiced] segments are *necessarily* [–spread glottis]: a (widely) spread glottis cannot vibrate (this claim is challenged below, however). The full set of implications, given in formula (1), can be depicted by adding a circle specified as [–spread glottis] outside the circle specified as [voiced], as seen in Figure 2 (from Basbøll, 2005:195). The outermost *circle ring* of Figure 2 – or more technically: the outermost *annulus* in the mathematical (geometric) sense – represents what is outside the *circle* with [–spread glottis] segments. This outermost circle ring of Figure 2, with no explicit specification, contains the complementary set to the set of [–spread glottis] segments, viz. the segments that are *not* [–spread glottis], i.e. they are [+spread glottis].

The full sequence of the sonority-hierarchical features in the Sonority Syllable Model depicted in Figure 2 is thus the one given in formula 1. But that *presupposes* that *[+voiced, +spread glottis] be an excluded type of segment.* Here, it is relevant to consider the so-called voiced aspirated plosives (oral stops) in many languages, as in e.g., Hindi, and also breathy or murmured vowels. I shall take the position here that [+voiced, +spread glottis] is not an excluded segment type. This means that

[+voiced] and [–spread glottis] cannot be part of the same version of the Sonority Syllable Model. I then take [spread glottis] to be a more important sonority-hierarchical feature than [voiced], in fact as important as [vocoid] and [sonorant]; there is no doubt in my mind that the implication below under formula 3 does hold:

(3) [+sonorant] *implies* [–spread glottis]

If breathy or murmured vowels are classified as [+spread glottis], they cannot be [+sonorant] in the definition used here (formula 3). I do not deny, however, that they have some characteristics of sonorants, but they are atypically so (with a less sharp formant structure, etc.), and I would not classify them as vocoids. In the remainder of this paper, I shall consider two versions of the Sonority Syllable Model, one with [+voiced], the other with [–spread glottis], the latter one being *preferred*, at least for languages with aspirated stops.

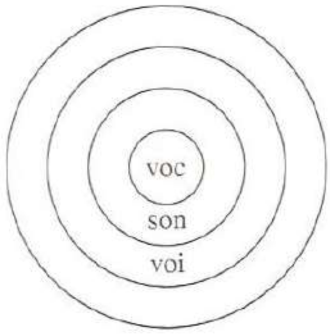


Figure 1. (from Basbøll, 2005:182). 4 (concentric) Euler's circles representing the implications [+vocoid] > [+sonorant] > [+voiced]. The outermost circle ring represents the complementary set to the set of [+voiced] segments, viz. those that are [–voiced]

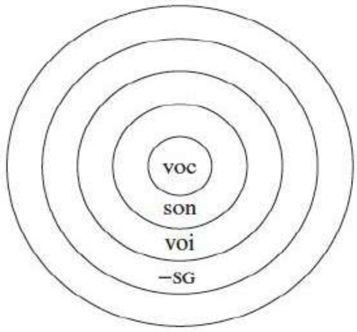


Figure 2 (from Basbøll, 2005:195). 5 (concentric) Euler's circles representing the implications [+vocoid] > [+sonorant] > [+voiced] > [–spread glottis]. The outermost circle ring represents the complementary set to the set of [–spread glottis] segments, viz. those that are [+spread glottis]

The Sonority Syllable Model: introduction of time, resulting in order classes

So far, no reference to *order*, *time*, *syllables*, or the like has been indicated: the model is purely *static* and just concerns the logical relation between segment types. But here comes the crucial point of the modeling, viz. the introduction of *time*. First, for simplicity's sake, I introduce time into Figure 1, thereby producing Figure 3 (from Basbøll, 2005:184):

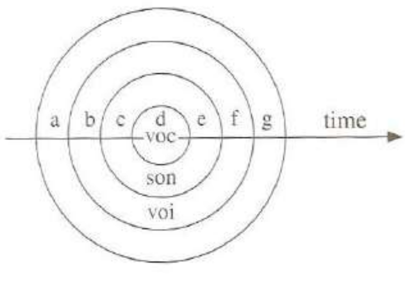


Figure 3 (from Basbøll, 2005:184). This figure is the same as fig. 1 but *with the dimension of time introduced* and is thus a *syllable model.* The letters a – b – c – d – e – f – g represent *order classes*

Figure 3 is a notational variant of a set of order classes, rendered in Figure 4 (from Basbøll, 2005:185). Recall that [–sonorant] and [+obstruent] are, by definition, identical. The seven letters: a – b – c – d – e – f – g of Figure 3 are represented by boxes with distinctive features in Figure 4.

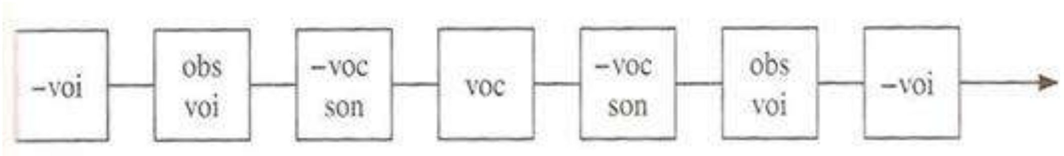


Figure 4 (from Basbøll, 2005:185). This is a notational variant of Figure 3, where the order classes are represented by boxes with distinctive features (thus representing *segment types*, more precisely *sonority-types*)

The Sonority Syllable Model is *unique*, I think, in its *non-circular and non-inductive* foundation, residing in general phonetic and phonological categorizations, including the *definition* of [vocoid]. The approach with the *Sonority Syllable Model* can be applied to all languages, even though not all consonant clusters in all languages obey all restrictions of the maximal model (see Basbøll, 2005:173-247, especially 193-201, on how languages with complex phonotactics can be described within this model). As said, the crucial point here is that the *Sonority Syllable Model* avoids the danger of circularity inherent in other models of sonority or strength hierarchy that either are generalizations of observed phonotactic patterns, i.e. the very kind of data they are supposed to explain, or build upon unfalsifiable innateness postulates.

The Sonority Syllable Model is based upon phonological segments classified with respect to specific general phonetic dimensions, viz. – in the preferred model – the four most important *sonority types*, defined by *sonority-hierarchical features* only (mentioned from the center outwards), i.e. [+vocoid], [–vocoid, +sonorant], [–sonorant, –spread glottis], [+spread glottis]. But the distinction between one and two identical segments must be found outside sonority, viz. in *prosody* (not least tonal patterns) and in *substitutability* (which positions can be established by a minimal pair test, for example?). A number of relevant such examples in Danish are discussed by Rischel (2003), e.g. *far* 'daddy', *fare* 'danger', *farere* 'travelers', all pronounced with [f] plus a single vowel quality [α] 'counting as', and perceived as, by Danish listeners, one, two, or three syllables, respectively.

It follows from the (preferred) Sonority Syllable Model that segments with *spread glottis* occupy the marginal position in the word. This agrees with the fact that an isolated word begins and ends with resting position (breathing). For instance, the model predicts that an initial plosive before /s/ (cf. Greek *ps-*), or a final plosive after /s/, will have a spread glottis; this is *testable* by direct observation of the glottis (cf. the end of section 4). Furthermore, there is a particular conclusion to be drawn about the 'mirror image'-issue:

The Sonority Syllable Model predicts that sequences specified in terms of *sonority types* exhibit *mirror image* structure initially and finally. This is generally true. But it does not follow from the Sonority Syllable Model that sequences of equal sonority are *mirror image*-like. On the contrary, they tend strongly to have the *same order initially and finally*, according to independent principles, including morphology, cf. clusters resulting from schwa-drop in the French prefix *r(e)-* (see Basbøll & Lambertsen, 2014). The (preferred) Sonority Syllable Model predicts that marginal segments in absolute initial and final position have [+spread glottis] e.g., *st-*, *ts-*, *-st*, *-ts*; in the same manner, presumably, we find final devoicing, e.g. in French *-fl* (*gifle* 'slap') (cf. Basbøll & Lambertsen, 2014).

Major classes derived from the model (3)

Several proposals have been presented (Basbøll, 2001, 2005:117-127) to depict major classes in a 2-dimensional space, by combining the *sonority-hierarchical features* and *[stop]*, but no entirely satisfactory solution was arrived at. Here, I propose what is in my view the optimal analysis departing from the *Sonority Syllable Model*, still including [stop], but omitting [voiced]. I start with the *minimal model*, operating with *5 Major Classes* only. The point of departure is the definition given in formula 2: building upon the definition of vocoids *only*, one can derive the 5 Major Classes in Table 1 (see Basbøll, 2001:89) by means of only three features: [vocoid], [sonorant] and [stop]. The remaining three logical possibilities (2^3 minus 5) from these three binary features are excluded by the very definition; this is a major advantage according to Occam's razor principle.

Table 1 (from Basbøll, 2001:89). Major Classes defined by the features [vocoid], [sonorant] and [stop]. Redundant feature values, by the definition of [vocoid], are parenthesized

	V	L	N	F	P	*	*	*
vocoid	+	–	(–)	(–)	(–)	+	+	+
sonorant	(+)	+	+	–	–	+	–	–
stop	(–)	–	+	–	+	+	–	+

Notice that the classes (segment types) L and N in Table 1 encompass only sonorant members (even though voiceless nasals and laterals do occur) that are those members that occupy a well-defined position in sonority hierarchies. The category ‘liquids’ is particularly ill-defined in this respect, cf. that ‘l-sounds’ and ‘r-sounds’ can be widely different with respect to ‘sonority’. One should therefore be skeptical of claims that ‘liquids’ constitute a natural class, phonetically and phonologically.

Modelising Major Classes in a two dimensional plane

Figure 5 (from Basbøll, 2001:91) shows that the three features [vocoid], [sonorant] and [stop] define five possible areas for Major Class segment types:

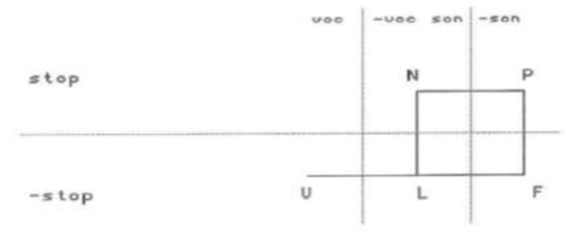


Figure 5 (from Basbøll, 2001:91). The 5 Major Classes of Table 1 depicted in a two-dimensional figure with [+vocoid] segments in the left column, [–vocoid, +sonorant] segments in the middle column, and [–sonorant] segments (obstruents) in the right column. [+stop] segments occur in the upper row and [–stop] segments in the lower row

Sonority-hierarchical features are [vocoid] and [sonorant] as employed here. *Sonority-hierarchical features are horizontal in the figure*, other features – here merely [stop] – are vertical in the figure. *The point of departure for the model is Vocoid in the bottom left corner. The figure is fully determined by the principles stated.*

A notational variant of Figure 5 is given in table 2, which is perhaps easier to read, and which makes redundancy of sonority-hierarchical features – from the definition of [vocoid] – explicit by parenthesizing redundant feature values; furthermore, [–stop] is redundant for V.

Table 2. A notational variant of Figure 5. Redundant feature values are parenthesized; also [–stop] is redundant for V (all redundancies follow from the definition of [vocoid])

	+ vocoid	–vocoid	(– vocoid)
	(+ sonorant)	+ sonorant	– sonorant
+ stop		**N**	**P**
– stop	**V**	**L**	**F**

Measuring distances between Major Classes

It makes sense to measure distances between the Major Classes defined here (see Figure 5 and Table 2) in two different ways, as illustrated in Table 3 (from Basbøll, 2001:91):

in ***sonority-steps strictly speaking*** (the horizontal dimension in Figure 5 and Table 2); or

in ***distance from V***, since nasals and plosives – which are [stop] – are clearly more distant from V than L or F respectively, since V is [–stop] by definition.

The distances are measured in a 'binary' way, i.e. a Major Class (segment type) is either in a particular field, or it is not, *tertium non datur*. The *sum of these two measures of distance* seems to capture an intuitive notion of 'strength', as often quoted in the literature, as well as of 'sonority', the values of which are inversely proportional to those of 'strength', see Table 3 (from Basbøll, 2001:91).

Table 3 (from Basbøll, 2001:91). Calculations of *sonority-steps* and *distance from V*, and of their sum for the 5 Major Classes

Major Class in the figure	V	L	N	F	P
sonority step	0	1	1	2	2
distance from V	0	1	2	2	3
sum of these	0	2	3	4	5

The final proposal: Seven Major Classes defined by the Sonority Syllable Model (4)

Departing from Table 2, which is a notational variant of Figure 5, we can add the sonority-hierachical feature [–spread glottis], thereby ending up with 7 Major Classes as illustrated in Table 4 (the notational conventions are the same as in Table 2).

Table 4. 7 Major Classes in 4 columns (with [+vocoid], [–vocoid, +sonorant], [–sonorant, –spread glottis] and [+spread glottis], respectively). Same conventions as for Table 2

	+ vocoid	– vocoid	(– vocoid)	(– vocoid)
	(+ sonorant)	+ sonorant	– sonorant	(– sonorant)
	(– spread glottis)	(– spread glottis)	– spread glottis	+ spread glottis
+ stop		**N**	**P unasp**	**P asp**
– stop	**V**	**L**	**F voiced**	**F voiceless**

When the seven Major Classes of table 4 are represented in terms of *sonority-steps* and *distance from V*, the results are something which in my opinion is the phonologically optimal table, viz. table 5, with [–spread glottis] in addition to [+vocoid] and [+sonorant]:

Table 5. Calculations of *sonority-steps* and *distance from V*, and of the sum of these, for the 7 Major Classes of table 4. Same conventions as for Table 3

	V	**L**	**N**	**Fvoi**	**Punasp**	**Fvless**	**Pasp**
son-d	0	1	1	2	2	3	3
V-dist	0	1	2	2	3	3	4
SUM	**0**	**2**	**3**	**4**	**5**	**6**	**7**

This yields an attractive strength or sonority hierarchy which is *not circular*.

Particularly for *obstruents*, this hierarchy presents some interesting predictions:

(a) *initially*, /s/ plus unaspirated plosive is *predicted*;

(b) *finally*, unaspirated plosive plus /s/ is *predicted*;

(c) if a plosive is more marginal than /s/ (e.g., *ps-*, *-sp*), it is *predicted* that the plosive is [+spread glottis]. Phonetically, anything else would be very surprising, and notice that the glottis position can be observed, so the proposal can be tested (see e.g., Löfqvist & Yoshioka (1980) for glottographic evidence).

Conclusion (5)

1) The set of 7 Major Classes derived from the sonority-hierarchical features [vocoid], [sonorant] and [spread glottis] *from the Sonority Syllable Model*, supplemented by the non-sonority-hierarchical feature [stop] – that is part of the definition of [vocoid], the point of departure of the whole approach – is the *preferable sonority hierarchy* and, in particular, *non-circular*.
2) For languages having a pure voice distinction in obstruents, but no aspirated stops, the inclusion of the feature [voiced] should be considered (instead of [–spread glottis]).
3) Phonotactic distinctions of a non-hierarchical nature (e.g., involving Place of articulation) should not be overlooked.

Acknowledgements

I am indebted to Nina Grønnum and an anonymous reviewer for many useful comments on the ms.

References

Basbøll, H. (1973). Notes on Danish consonant combinations. *Annual Report of the Institute of Phonetics, University of Copenhagen*, *7*, 103-142.

Basbøll, H. (2001). What can be derived from just three binary features: Occam's razor and major classes for phonotactics. *Travaux du Cercle Linguistique de Copenhague*, *21*, 74-99.

Basbøll, H. (2005). *The phonology of Danish* (Series: Phonology of the World's Languages). Oxford, UK: Oxford University Press.

Basbøll, H., & Lambertsen, C. (2014). La syllabe de sonorité rencontre le mot français : Aspects de morphonotaxe. In J. Durand, G. Kristoffersen, B. Laks, & J. Peuvergne (eds.), *La phonologie du français : Normes, périphéries, modélisation* (pp. 83-99). Paris: Presses Universitaires de Paris Ouest.

Ladefoged, P. (1971). *Preliminaries to linguistic phonetics.* Chicago, IL: University of Chicago Press.

Löfqvist, A., & Yoshioka, H. (1980). Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America*, *68*, 792-801.

Ohala, J. J. (1992). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. *Papers from the Parasession on the Syllable* (pp. 319-338). Chicago, Il: Chicago Linguistic Society.

Ohala, J. J. (2008). The emergent syllable. In B. L. Davis & K. Zajdo (eds.), *The syllable in speech production* (pp. 179-186). Erlbaum, UK: Francis & Taylor.

Ohala, J. J., & Kawasaki-Fukumori, H. (1997). Explaining segmental sequential constraints. In S. Eliasson & E. H. Jahr (eds.), *Language and its ecology: Essays in memory of Einar Haugen* (pp. 343-365). Berlin: Mouton de Gruyter.

Pike, K. L. (1943). *Phonetics: A critical analysis of phonetic theory and a technique for the practical description of sounds.* Ann Arbor, MA: University of Michigan Press.

Rischel, J. (2003). The Danish syllable as a national heritage. In H. Galberg Jacobsen, D. Bleses, T. O. Madsen, & P. Thomsen (eds.), *Take Danish – for instance. Linguistic studies in honour of Hans Basbøll* (pp. 273-281). Odense: University Press of Southern Denmark.

To cite this article: Basbøll, H. (2017). The Sonority Syllable Model reconsidered: Two challenges and their solution. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 71-77). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Cross-linguistic syntactic priming in Chinese-English bilinguals

Liang Chen[1], Xiangyu Jiang[2], Qin Zhou[3]
chen@uga.edu

[1]University of Georgia, [2]Harbin Institute of Technology at Weihai,
[3]Central South University

Abstract. Cross-linguistic syntactic priming (CLSP) refers to the phenomenon that the use of a particular construction (e.g., passives) in one language primes subsequent use of that construction in the other language of a bilingual speaker. Previous research showed no evidence of CLSP in the passive construction of language pairs, such as German and English that differ in the relative surface word of the main verb and the *by*-phrase (Loebell & Bock, 2003). However, this construction can be primed cross-linguistically between language pairs, such as English and Spanish that show similar surface word order (Hartsuiker, Pickering, & Veltkamp, 2004). More recently, however, the degree of CLSP has been found to be influenced by the relative proficiency of bilinguals' two structurally similar languages (e.g., Kim & McDonough, 2008; Shin & Christianson, 2009). Consequently, the effect of structural similarity remains unclear and may be merely an artifact of language proficiency. The present study was designed to address this question with bilingual speakers of Chinese and English who languages differ in the surface word order of passives. 50 Chinese learners of English at two proficiency levels (HIGH vs LOW) participated in a picture-description task. The task consisted of describing pictures in English after listening to prime sentences in Chinese or English. Results revealed a CLSP effect and that the HIGH proficiency group was more likely to use a passive in English after passive primes than the LOW proficiency group. We conclude that CLSP doesn't require the surface word order of the passive construction to be the same in both languages, but the effect of structural similarity may be modulated by the proficiency level of the bilingual participants.

Keywords: cross-linguistic syntactic priming, Chinese-English bilinguals, language proficiency, structural similarity

Introduction

Syntactic priming is the tendency for a speaker to reuse the same structural pattern s/he has just heard or produced (e.g., Bock, 1986; Potter & Lombardi, 1998). Bock (1986), for example, asked participants to describe pictures in English after repeating an auditory stimulus of prime sentences presented in English. These sentences were semantically unrelated to target sentences. In one manipulation, the primes were either in prepositional-object (PO) construction, such as *The girl handed a paintbrush to the man,* or in the alternative double-object phrase structure (DO), such as *The girl handed the man a paintbrush*. Bock found that having repeated aloud a sentence, such as *The rock star sold some cocaine to an undercover agent*, participants were more likely to describe a picture with the same prepositional-object (PO) construction, rather than with a double-object phrase structure (DO).

Recently, an increasing number of studies have examined syntactic priming in bilingual individuals (e.g., Schoonbaert, Hartsuiker, & Pickering, 2007; Loebell & Bock, 2003; Meijier & Fox Tree, 2003; Song & Do, 2016). Cross-linguistic syntactic priming (CLSP) refers to the phenomenon that the use of a particular structure (e.g., passives) in one language facilitates or primes the subsequent use of that structure in the other language of a bilingual speaker. An unresolved issue of CLSP is whether it is manifested when the two languages of the bilingual differ in surface word orders of the relevant constructions. Previous research has produced mixed results regarding whether surface word order similarity between the two languages is a prerequisite for CLSP. For example, in bilingual speakers of English and Spanish, CLSP was observed with similar surface word order of passives (Hartsuiker, Pickering, & Veltkamp, 2004), and no CLSP for German and English passives that differ in the relative order of the main verb and the *by*-phrase (Loebell & Bock, 2003). While English and Spanish passives (e.g., 1a and 1b) share similar structures with agent-marking by-phrases following the verb in English, German passives can be formed with the main verb in the final position of the sentence, as in (1c).

(1) a. The floors are cleaned daily by the janitor. [English]
b. *Los pisos son limpiados diariamente por el conserje*. [Spanish]
'the floors are cleaned daily by the janitor'
c. *Die Bőden warden täglich von dem Hausmeister gereinigt*. [German]
(literally: 'the floors are daily by the janitor cleaned.')
d. *Dìbăn mĕitiān dōu bèi qīngjiégōng dăsăo de hen gānjìng*. [Chinese]
floor everyday all by janitor sweep DE very clean
(literally: "The floors daily by the janitor cleaned.")

By contrast, Chen, Jia, Wang, Dunlap, and Shin (2013) found priming of the passive structure between Chinese (L1) and English (L2), despite the different word orders of passives in the two languages. Chinese is similar to German, except that the auxiliary *be* is not used in Chinese. Specifically, as shown in (1d), the phrase denoting the agent (i.e., the *bèi*-phrase) in Chinese is placed before the verb, whereas the by-phrase in English is placed in sentence-final position after the verb. Despite this word order difference, Chen *et al*. (2013) observed passive priming not only from Chinese to English but also from English to Chinese, by using two different experimental paradigms (i.e., a picture description and a confederate-scripting paradigm).

Another controversial issue is the role of language proficiency in CLSP. Some studies suggest that the degree of CLSP is influenced by the relative proficiency of bilinguals' two structurally similar languages. Through a picture description activity participated by Korean learners of English from three proficiency levels, Kim and McDonough (2008) found that learners were primed to produce more passives when the same verbs had occurred in the researcher's passives, but low proficiency learners relied more on individual lexical items than advanced learners. Similarly, Shin and Christianson (2009) showed that CLSP from English (L2) to Korean (L1) with dative structures is limited to proficient L2 speakers (see also McDonough, 2006; McDonough & Fulga, 2015). Bernolet, Hartsuiker, and Pickering (2013) examined priming of possessive structures in Dutch-English bilinguals who spoke Dutch as L1 and English as L2. They assessed participants' proficiency in each language via self-rating and found that priming from both L1-to-L2 and L2-to-L2 increased in magnitude as proficiency in L2 increased. Other studies, by contrast, indicate that although passive structures prime across languages, language proficiency does not affect priming. Cooperson (2013) for example, asked sixty-eight Spanish-English bilingual adults to complete L1-to-L2 and L2-to-L2 passive priming tasks and measures of language experience and proficiency. Results revealed CLSP in both directions, and more importantly the variables of age of acquisition, current use, and proficiency were not found to affect priming.

In summary, previous studies have produced conflicting results regarding the role of bilinguals' proficiency levels and surface word order similarities in CLSP. In addition, no study, to our best knowledge, has considered both surface word order similarity and language proficiency in CLSP. The present study was designed to address this research gap through an examination of CLSP of passive structures in Chinese-English bilingual speakers. Specifically, we address the following three research questions.

(2) a. will Chinese-English bilingual speakers exhibit syntactic priming crosslinguistically and within English?
b. will there be greater syntactic priming within-language than between-languages?
c. if the crosslinguistic priming occurs, will Chinese-English bilingual speakers' proficiency levels play a role in the syntactic priming tasks?

Method

Participants

The participants were 50 Chinese university students recruited from a university in Northern China. All were native speakers of Chinese and were learning English as their L2. The learners had a mean age of 21.3 years (SD=2.27) and had studied English previously for more than ten years. The participants were divided into two groups. The high proficiency group included 28 English majors who had undergone intensive English courses in college for two and a half years at the time of data collection. The low proficiency group included twenty-two non-English majors who had been university students for half a year.

Materials

There were two sets of 48 cards, 96 cards in total, each portraying an action. For the participants' cards, all 16 target pictures contain an agent carrying out an action, and the patient undergoing an action. The agents were inanimate in all 16 cards, but patients were inanimate in 8 cards and animate in the other 8 cards. Agents were always shown on the right side of the card. The remaining 32 filler cards can only be described with intransitive sentence structures (e.g., Figure 2). On the participants' cards, a verb was written at the bottom of each card. Participants were required to use the given verb in their response (e.g., Figure 1). For the researcher's set, only prime sentences, not pictures were labeled on the card. The 16 prime sentences described 8 events. Each event was described once by active structure and once by passive structure. Each prime sentence was preceded by 2 filler sentences. The researcher's set was paired with the participant's set, so that each prime sentence could be immediately followed by a target picture. Furthermore, prime sentences were designed and organized in such an order that half of the prime sentences had the same animacy as the corresponding target picture, including both the agent and the patient. The other half had an agent and patient with the opposite animacy. The verbs used in the prime and target were never translation equivalents or repeated.

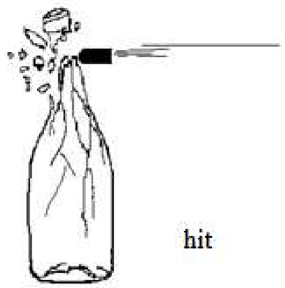


Figure 1. A target picture example (a bullet is hitting a bottle. Hartsuiker et al., 2004)

Figure 2. A filler picture example (A girl is coughing)

Procedure

The participants were randomly placed in either the within-language (English-to-English) priming condition or the crosslinguistic (Chinese-to-English) priming condition. Each participant met individually with a researcher for approximately 15 minutes. The researcher explained to the participant that they were going to describe pictures in turn. Then, a short practice session was given to familiarize the participant with the process. After the practice session in the crosslinguistic priming condition, the researcher and the participant took turns to describe their pictures. However, the researcher always preceded the participant in the picture description task to ensure that the passive priming sentences were produced before the participant described his/her pictures. 12 high proficiency participants and 15 low proficiency participants were included in the crosslinguistic priming condition. In the within-English priming condition, because the experiment was conducted in China, the participant heard audio recordings of picture descriptions recorded by a female American speaking standard American English. From time to time, the researcher would ask the participant to repeat what they have just heard to make sure the participant was listening to the recording carefully. Whenever the participant encountered unknown English words, s/he could turn to the researcher for help. 10

high proficiency participants and 13 low proficiency participants were included in the within-English priming condition. The entire experiment session was audio-recorded and conducted in a quiet room at the university.

Scoring

The participants' responses were transcribed and coded as passive, active, or other. A response was scored as a passive if it contained a patient, an auxiliary verb, a verb past participle, regardless of whether the agent was expressed in a *by*-phrase. A special type of sentences (20 in total, 2.7%) is also regarded as passive. For example, **the man is striking by the lightning*. This type of sentence contains all features of a passive, except the wrong use of past participle, which may be caused by the fact that Chinese does not contain the corresponding past participle morpheme. Sentences with an agent serving as the subject and the patient as the object were regarded as actives. Other sentences include: auxiliary verbs were omitted, agents occupied the subject position in passive sentences, and objects were neglected in active sentences. Grammatical errors, such as subject-verb agreement, verb tense, and verb past participle forms were not considered in this study. A bilingual Chinese-English speaker transcribed and coded all the responses in the study. A second trained rater transcribed and coded 20% of the responses separately. The two raters then checked whether their transcriptions matched. Initial agreement of the raters' transcriptions was 94.6%. They re-listened to the recording together until consensus was reached. The raters then checked whether their coding of the sentences (active or passive) matched. Initial agreement of the coding was 96.5% of responses. For the responses with disagreement, the first and second authors decided the correct coding.

Results

There were a total of 752 target descriptions with 230 active (30.6%), 437 passive (58.1%), and 85 other (11.3%). The most frequent types of 'other' descriptions were passive sentences without auxiliary verbs (e.g., **the woman will run over by train*), and active sentences without patients (e.g., *the thunder is striking*). We computed the relative proportions of active and passive target responses in each of the priming conditions. This measure represented the number of active or passive target responses divided by the sum of active target responses plus passive target responses. Table 1 presents these proportions by bilinguals with different levels of English proficiency in the within-language (English-to-English) and crosslinguistic (Chinese-to-English) priming conditions.

Table 1. Mean proportions of active (A) and passive (P) target responses as a function of prime structure, English proficiency and priming conditions

	Prime	**Target**	
		A	P
Low Proficiency Group			
Chinese-to-English	A	42.92%	57.08%
	P	38.63%	61.37%
English-to-English	A	53.79%	46.21%
	P	38.57%	61.43%
High Proficiency Group			
Chinese-to-English	A	25.95%	74.05%
	P	13.68%	86.32%
English-to-English	A	44.58%	55.42%
	P	29.29%	70.71%

Structural priming is demonstrated when the participant responses (i.e. target structures) match the structure of the prime on more than 50% of the trials (Loebell & Bock, 2003; Coyle & Kaschak, 2012). As Loebell and Bock (2003:805) note 'proportions greater than 0.50 indicate increased use of a particular form after priming; correspondingly, proportions less than 0.50 indicate decreased use of the form after priming'. Table 1 shows the passive responses after passive primes range from 61% to 86%, and our participants were 28.15% more likely to produce a passive target response than an active target response following either an active or passive prime. These percentages of passive responses are higher than what's typically reported in the literature, except for Flett's (2003) study of syntactic priming of passives in L1 and L2 Spanish. Flett found that native speakers of Spanish produced 28% passive, while advanced English learners of Spanish as an L2 produced 73% passives, and the intermediate English learners of L2 Spanish produced 58% passives following a passive prime. While the high proportions of passive responses remain puzzling, these results suggest that our study produced the standard syntactic priming effect for passive structures, regardless of English proficiency and priming condition.

Since our participants produced more passive structures (except in the English-to-English condition with low proficiency group) even after active primes, contrary to our expectation, we focused on the analysis of passive target responses after passive primes. The mean proportions of passive target responses after passive primes (see Figure 3) were subject to a two-way ANOVA (English proficiency x priming condition). The main effect for English proficiency is significant ($F_1(1,49)=44.1$, $p<.001$; $F_2(1,15)=10.63$, $p<.005$), with the high proficiency group (M=78.5%, SD=0.36) being 17% more likely to produce a passive target response after a passive prime than the low proficiency group (M=61.4%, SD=0.23). It was also found that there was a tendency for the priming condition to influence the likelihood of passive target response after a passive prime, and that participants were 8% more likely to produce a passive target response after a passive prime in the Chinese-to-English priming condition (M=73.8%, SD=0.28) than in the English-to-English priming condition (M=66.1%, SD=0.33). This difference, due to the priming condition, didn't reach significance at <0.05 level ($F_1(1,49)=3.50$, $p=0.07$; $F_2(1,15)= 3.52$, $p=.08$). A significant interaction between English proficiency and priming condition was found, $F_1(1,49)=10.94$, $p<.01$; $F_2(1,15)=13.6$, $p<.005$. Inspection of the means suggested that the expected effect of priming condition occurred only for the high proficiency group. That is, only the high proficiency group participants were more likely to produce passive target responses in the Chinese-to-English condition (M=86.3%, SD=0.56) than in the English-to-English condition (M=61.3%, SD=0.45), $t(49)=3.12$, $p < 0.01$.

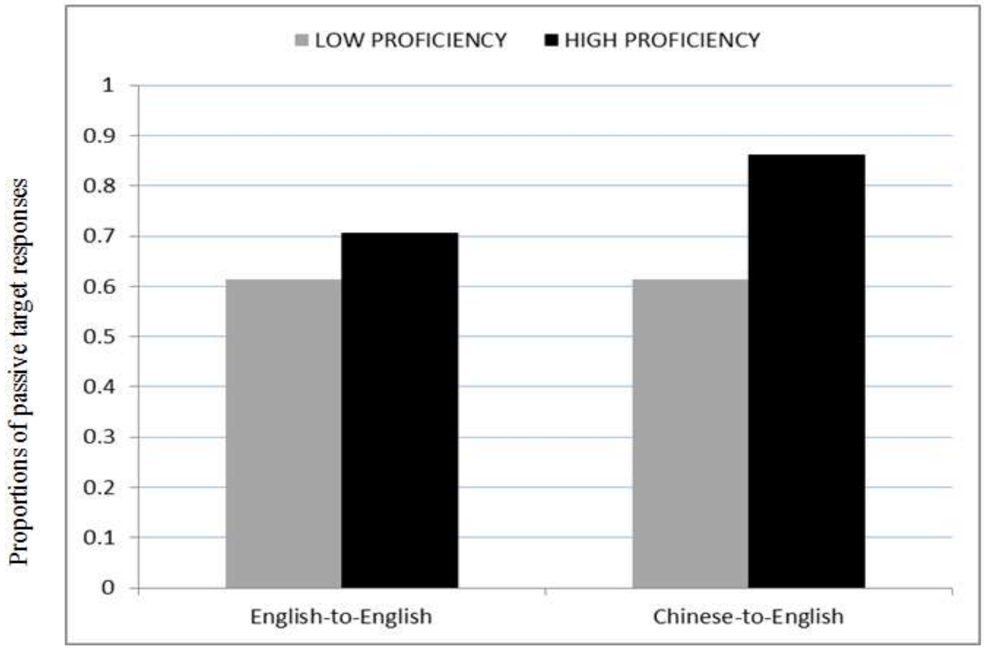


Figure 3. Mean proportions of passive target responses after passive primes as a function of English proficiency and priming condition

Discussion and conclusions

The present study revealed three important findings. First, it showed priming of passive structures in both the within-English and the crosslinguistic conditions in Chinse-English bilinguals. These results are consistent with those of Chen *et al*. (2013), and provide further evidence that CLSP of passives can occur independently of structural similarities.

Secondly, language proficiency was found to affect the degree of L1-to-L2 priming of passive structures in Chinese-English bilingual participants. These results are consistent with previous studies, such as Bernolet *et al*. (2013) and Shin (2010), who found that L1-to-L2 priming became stronger with increasing L2 proficiency. While previous studies have demonstrated the effect of language proficiency on the priming between structurally similar languages, the present study provides evidence for the effect of language proficiency on CLSP, even when the two languages of the bilinguals differ in the surface word order of the passive structures.

Finally, this study showed that language proficiency affected the degree of priming on within-language priming in the Chinese-English bilingual participants. These results are consistent with those obtained by Flett (2003), who examined the effects of language experience and proficiency on within-language priming of Spanish passive structures in 12 native speakers of Spanish, 12 intermediate and 12 advanced English-speaking learners of Spanish as a second language. The level of these L2 speakers was assessed using a cloze test. Flett (2003) found that priming was stronger in L2 speakers than native speakers, and it was stronger in L2 speakers of advanced proficiency than L2 speakers of intermediate proficiency. While Flett (2003) provides evidence for the effect of language proficiency on syntactic priming between English and Spanish passives that share similar word order, the present study suggests the effect of language proficiency on syntactic priming, even when the two languages of the bilinguals differ in the word order of the passive structures.

These results lead us to conclude that CLSP doesn’t require the surface word order of the passive construction to be the same in both languages. Instead, the effect of structural similarity may be modulated by the proficiency levels of the bilingual participants.

Acknowledgements

We acknowledge the following financial support: to Xiangyu Jiang from the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Youth Foundation Project, Grant No. 17YJC740032) and from the Teaching and Research Program of Harbin Institute of Technology at Weihai (Youth Special Program, Grant No. ITIA 10002010); to Qin Zhou from Hunan Province Education Scientific Planning Committee Research Project (Grant NO. XJK015BGD088).

References

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2007). Shared syntactic representations in bilinguals: Evidence for the role of word-order repetition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 931-949.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). Effects of phonological feedback on the selection of syntax: Evidence from between-language syntactic priming. *Bilingualism: Language and Cognition*, *15*, 503-516.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355-387.

Chen, B., Jia, Y., Wang, Z., Dunlap, S., & Shin, J.-A. (2013). Is word-order similarity necessary for cross-linguistic structural priming? *Second Language Research*, *29*(4), 375-389.

Cooperson, S. J. (2013). *Priming across languages in Spanish-English bilinguals*. Unpublished PhD Dissertation, The University of Texas at Austin.

Coyle, J. M., & Kaschak, M. P. (2012). Female fertility affects men's linguistic choices. *PLoS ONE*, *7*(2), e27971.

Flett S. (2003). Syntactic priming in L1 and L2 Spanish. Paper presented at the 2003 Linguistics Postgraduate Conference, University of Edinburgh.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science*, *15*, 409-414.

Kim, Y., & McDonough, K. (2008). Learners' production of passives during syntactic priming activities. *Applied Linguistics*, *29*, 149-154.

Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics*, *41*, 119-135.

McDonough, K. (2006). Interaction and syntactic priming: English L2 speakers' production of dative constructions. *Studies in Second Language Acquisition*, *28*, 179-207.

McDonough, K., & Fulga, A. (2015). The detection and primed production of novel constructions. *Language Learning*, *65*, 326-357.

Meijier, P. J. A., & Fox Tree, J. E. (2003). Building syntactic structures in speaking: A bilingual exploration. *Experimental Psychology*, *50*, 184-195.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentence. *Journal of Memory and Language*, *38*, 265-282.

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language*, *56*, 153-171.

Shin, J, -A., & Christianson, K. (2009). Syntactic processing in Korean-English bilingual production: Evidence from cross-linguistic syntactic priming. *Cognition*, *112*, 175-180.

Song, Y., & Do, Y. (2016). Cross-linguistic structural priming in bilinguals: Priming of the subject-to-object raising construction between English and Korean. *Bilingualism: Language and Cognition*. https://doi:10.1017/S1366728916001152

To cite this article: Chen, L., Jiang, X., & Zhou, Q. (2017). Cross-linguistic syntactic priming in Chinese-English bilinguals. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 78-84). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Speechreading performance of Chinese adolescents with and without hearing impairment

Liang Chen[1], Jianghua Lei[2], Huina Gong[2]
chen@uga.edu, huajianglei2003@163.com, gonghn1991@126.com

[1]University of Georgia, [2]Central China Normal University

Abstract. Evidence regarding the effect of hearing status on speechreading skills is equivocal. A speechreading advantage has been reported for deaf or hard of hearing individuals (e.g., Tye-Murray, Hale, Spehar, Myerson, & Sommers, 2014) but also for individuals with normal hearing (Green, Green, & Holmes, 1981). Moreover, some studies have found no effect of hearing status on speechreading ability in English-speaking children (e.g., Kyle, Campbell, & MacSweeney, 2016). These studies have been carried out primarily on European/Western languages, especially English (Chen & Lei, 2017). Thus, the present research set out to investigate the effect of hearing status on speechreading skills in Chinese speakers with and without hearing impairment. Thirty seven severely deaf students with a mean pure-tone average of 93 dB HTL and twenty one hearing controls aged 16 were asked to complete tasks measuring their speechreading of simplex finals (monophthongs), complex finals, and initials (consonants) in Chinese. Both accuracy rate and response time data were collected. Results show that hearing individuals were in general more accurate in speechreading than their deaf controls, but the difference did not reach statistical significance ($F(1, 56)=0.733$, $p=0.396$). By contrast, deaf individuals were significantly faster at speechreading than their hearing controls ($F(1, 56)=24.13$, $p<.001$). For both groups, however, performance on speechreading simplex finals was faster and more accurate than complex finals which, in turn was better than initial consonants. The results of the present study show that speechreading skills in Chinese speakers are influenced by hearing status, the characteristics of sounds to be identified, as well as the measures used.

Keywords: speechreading, hearing status, deafness, Mandarin Chinese, visual speech perception

Introduction

Speechreading (visual-alone speech perception) contributes to speech processing, and the ability to speechread is often critical for effective spoken communication, particularly for individuals with hearing loss (Gagné, Charest, Le Monday, & Desbiens., 2006; Tye-Murray *et al*., 2014). Although the importance of visual speech perception via speechreading is well established, we are only just beginning to unravel its complexity. In particular, the effect of hearing status on the ability to speechread is poorly understood, and the evidence regarding differences in speechreading performance between children and adults with hearing loss (HL) and those with normal hearing (NH) is equivocal. Some studies reported that deaf or hard of hearing individuals have better speechreading ability in comparison to hearing controls (Auer & Bernstein, 2007; Bernstein, Demorest, & Tucker, 2000; Lyxell & Holmberg, 2000; Mohammed, Campbell, MacSweeney, Barry, & Coleman, 2006; Tye-Murray *et al*., 2014). Tye-Murray *et al*. (2014), for example, compared 40 children with NH (mean age =10;9, range=84-179 months) and 24 children with HL (mean age=10;10, range=91-177 months) in their performance on a speechreading battery which included four measures. They found that children with HL performed significantly better than those with NH on all four speechreading measures across the age range under consideration. On the other hand, Conrad (1977) compared a group of 486 hearing-impaired 15 to 16 year-olds who had hearing losses in excess of 64 dB HL with their hearing peers. The participants were shown between six and nine photographs and had to choose one of these to match the spoken stimulus. The results showed a slight, but non-significant, difference in favor of the hearing impaired. Elphick (1996) compared the speechreading performance of 57 hearing-impaired 15 to 16-year-olds with 30 hearing students in the same school year. The test stimuli consisted of two word lists, one from the AB isophonemic word list and the other from the New

Manchester picture test together with a sentence list from the Manchester speechreading (Lipreading) Test. The hearing-impaired group scored significantly better than the hearing group in all of the lists.

Other studies, however, found that individuals with normal hearing might have comparable or even superior speechreading skills in comparison to individuals with HL. Green *et al*. (1981) compared the performance of 22 deaf children (mean age= 6;3) to that of 22 normal hearing children (mean age= 5;9) on the diagnostic test of speechreading (Myklebust & Neyhus, 1970). Speechreading performance was assessed across word, phrase, and sentence stimuli. Even though the deaf children had been enrolled in an education program which emphasized the development of visual skills, results revealed significantly better performance by the normal hearing children on all stimuli. More recently, Kyle, Campbell, Mohammed, Coleman, and MacSweeney (2013) administered the test of child speechreading (ToCS) to 86 deaf and 91 hearing children between the ages of five and fourteen. ToCS measured child speechreading at 3 psycholinguistic levels: (a) words, (b) sentences, and (c) short stories. Kyle and her colleagues found a main effect of psycholinguistic level and a main effect of age on speechreading performance, but no effect of hearing status was found (see also Arnold & Köpsel, 1996; Conrad, 1977; Kyle *et al*., 2016).

One may have noticed that the studies reviewed in this introduction have reported an advantage for the deaf and hard of hearing participants across quite a large age range together for comparison. The lack of age-subgroup comparisons may have skewed the results. In addition, these studies have been carried out on European/Western languages, and primarily English (Chen & Lei, 2017). Thus, the present research set out to investigate the effect of hearing status on speechreading performance in Mandarin Chinese.

Method

Participants

The participants came from two groups, one with hearing impairment and the other without. The hearing-impaired participants consisted of thirty-seven students with prelingual sensorineural hearing loss (mean age= 17;1, SD=0.7). Unaided hearing sensitivity on the better ear, estimated by pure-tone average hearing threshold levels (HTLs) at 50, 1000, and 2000 Hz, averaged 93 dB (SD = 2.49). All of the hearing impaired participants were hearing aid users, and they were tested while wearing their devices. All of the participants were recruited from the same three schools for the deaf and hard of hearing in central China, where they were enrolled in their age appropriate grade. They were proficient sign language users, but there was an emphasis at school on spoken language training and the use of sign language was discouraged. The hearing control students included twenty-one participants (mean age= 16;6, SD=0.47). All of them were monolingual native Mandarin Chinese speakers, and were enrolled in schools around Wuhan city of Hubei Province in central China at the time of data collection.

It was confirmed, through independent-samples *t*-tests, that there were no differences in age between the hearing impaired group and the normal hearing group, $t(57)=2.6$, ns. All of the participants had normal or corrected-to-normal vision. No neurological, psychological, cognitive or language impairment was reported.

Materials

The speechreading material included test words for identification of initials (consonants) including /b/, /z/, /t/, / tʂʰ/, /tɕʰ/, /k/; the simplex finals (monophthongs) including /a/, /ɔ/, /ɤ/, /i/, /ʊ/, /y/, and the compound finals including /aɪ/, /ʊɔ/, /jɛ/, /an/, /ʊŋ/ and /ɑʊ/. Each target sound was assessed with eight words, four that contained the target sound and four that contained a sound with an identical place of articulation. A female speaker of standard Chinese (the common language) read the test words in a sound-attenuated room in front of a Sony TRV38E DV camcorder. The video recordings with sound were then edited so that the start and end frames of each token showed the female speaker from the shoulders up producing one of the test words. Three types of stimuli were prepared with

Adobe Authorware: visual-only (V), auditory-only (A), and audiovisual (AV) stimuli. However, speechreading performance was assessed with the V-type stimuli only in the present study (see Chen & Lei, 2017; Lei & Chen, 2015 for details).

Procedure

The procedure is the same as used in Lei and Chen (2015), and Chen and Lei (2017). Participants were tested individually. They sat in front of a computer and responded to test stimuli by pressing buttons. Instructions were given in sign language and in writing, and participants completed practice trials to demonstrate that they understood the instructions.

After seeing each word produced by the female speaker, participants viewed a response screen displaying the YES and NO buttons together with the question asking whether the word contains one of the target sounds. They made judgments as to whether the word produced contained a target sound by pressing either the YES button or the NO button. Participants were encouraged to guess if they were uncertain about a response. After each response, no feedback regarding the correctness of responses was provided, and participants were asked to click a ‘continue’ button to watch the next video stimulus. The presentation software recorded the accuracy rate and response time automatically. During the presentation, the experimenter monitored participants in order to make sure they were watching the screen. Each participant required approximately 20 minutes to complete the test.

Results

The dependent variables of interest consisted of mean accuracy rates and response times. These data are presented in Table 1 as a function of hearing status and test stimulus type. An analysis of variance (ANOVA) was performed on each dependent variable separately, with hearing status (Hearing vs. Deaf) and test stimuli type (initials, simplex finals, and compound finals) as between-subjects factors.

Table 1. Accuracy rate (% correct) and mean response time (RT in seconds) according to group and test stimuli types in severely deaf children and their hearing controls

Group	Initials		Compound Finals		Simplex Finals	
	Accuracy	RT	Accuracy	RT	Accuracy	RT
HEARING	0.538±0.119	3.70±1.75	0.626±0.112	3.03±1.01	0.70±0.148	3.02±0.97
DEAF	0.520±0.060	2.22±1.24	0.605±0.089	2.00±1.02	0.688±0.121	2.13±0.84

Accuracy rate

The mean accuracy rates are plotted in Figure 1. A 2 (hearing status) x 3 (test stimuli type) ANOVA was conducted to assess differences in speechreading accuracy. Neither the main effect of hearing status, $F(1, 52)=0.733$, p=0.396; nor the Hearing Status x Test Stimuli Type interaction, $F(2, 52)=0.43$, p=0.96, reached the level of statistical significance. Thus, the results do not provide evidence for the existence of a difference in identifying any of the three types of speech sounds between the participant groups. The ANOVA results, however, showed a significant main effect for stimulus type, $F(2, 52)=46.95$, $p<0.001$. Post hoc Scheffé tests conducted on the main effect of test stimulus type shows that deaf and hearing participants performed better on simplex finals than on compound finals ($p<0.001$) and initials ($p<0.001$) and, in turn, scored higher on compound finals than on initials ($p<0.001$).

Response time

The group means for response time are plotted in Figure 2. A separate 2 (hearing status) x 3 (test stimulus type) ANOVA was used to investigate differences in the time lapse prior to responding to the items of the speechreading task. Results revealed a significant main effect of hearing status, $F(1, 52)=24.13$, $p<0.001$. Deaf individuals were significantly faster at speechreading than their hearing controls. There was also a significant main effect of test stimuli type, $F(2, 52)=8.29$, $p<0.001$, but no significant interaction $F(2, 52)=1.18$, $p=0.31$. Scheffé post hoc tests showed that speechreading of initial consonants ($p<0.001$) showed longer response times than the speechreading of both simplex finals and compound finals. No significant differences in the speed of visual speech perception were found between the simplex finals and compound finals ($p<0.93$).

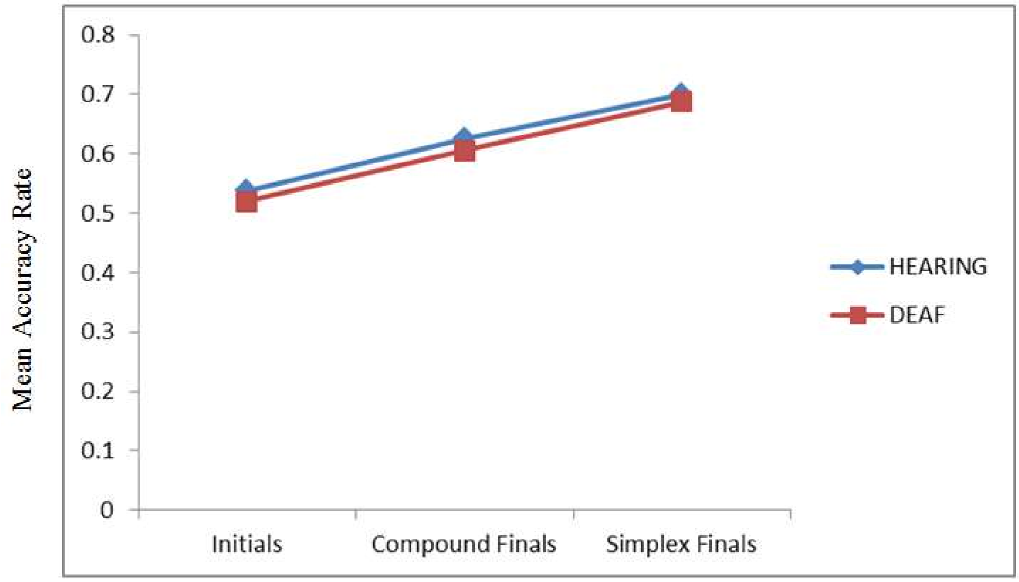


Figure 1. Mean speechreading accuracy rate as a function of hearing status and test stimuli type

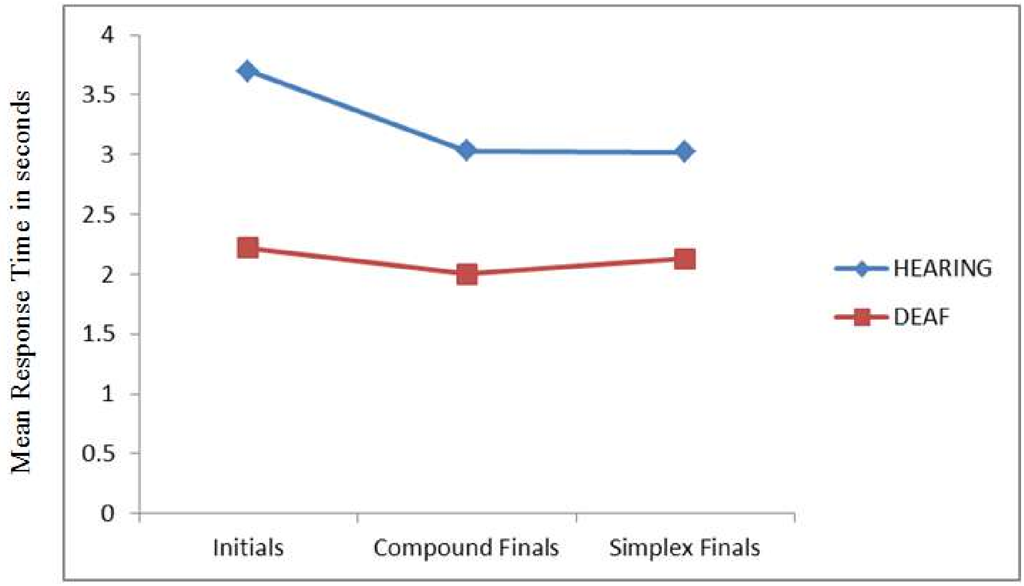


Figure 2. Mean response time as a function of hearing status and test stimuli type

Response time

The group means for response time are plotted in Figure 2. A separate 2 (hearing status) x 3 (test stimulus type) ANOVA was used to investigate differences in the time lapse prior to responding to the items of the speechreading task. Results revealed a significant main effect of hearing status, $F(1, 52)=24.13$, $p<0.001$. Deaf individuals were significantly faster at speechreading than their hearing controls. There was also a significant main effect of test stimuli type, $F(2, 52)=8.29$, $p<0.001$, but no significant interaction $F(2, 52)=1.18$, $p=0.31$. Scheffé post hoc tests showed that speechreading of initial consonants ($p<0.001$) showed longer response times than the speechreading of both simplex finals and compound finals. No significant differences in the speed of visual speech perception were found between the simplex finals and compound finals (p<0.93).

Discussion

The current study shows that hearing status has an effect on the speed, but not the accuracy of speechreading performance in Chinese sixteen-year-old adolescents. Specifically, deaf individuals were significantly faster at speechreading than their hearing controls. While the results of the present study still leave the question open whether the hearing impaired develop greater skills in speechreading than the hearing population, it is interesting to compare the study with two other studies that have compared 15 to 16 year-old hearing students with their hearing-impaired peers (Conrad, 1977; Elphick, 1996). Elphick (1996) pointed out that Conrad presented test stimuli which were spoken aloud to the hearing group and used masking to prevent auditory cues from hearing, but no specific procedure was described to ensure that every auditory cue was in fact eliminated. These factors might have enhanced the performance of the hearing group in Conrad's (1977) study, and if so would have masked the deaf and hard of hearing advantage in speechreading. The hearing-impaired students in Elphick (1996) were placed full time in either mainstream units for the hearing impaired or special schools for the hearing impaired. Elphick attributed the advantage of the hearing impaired in his study to a greater need for speechreading and the practice derived from it. Oliveira Soares, and Chiari (2014), in their study of speechreading in Brazilian Portuguese, also found a close correlation between speechreading therapy and speechreading performance, suggesting that there was changes in their participants' speechreading performance after a short period of practice. The participants from the present study were drawn from special schools for the hearing impaired, and thus had more speechreading practice than their hearing peers. It is reasonable to suggest that greater need and more practice might account for the faster speechreading performance of the hearing impaired participants. However, such an explanation would not contribute to our understanding of the lack of difference in speechreading accuracy in the present study.

Oliveira *et al*. (2014) found an advantage of the hearing impaired group on all measures except for speechreading sentences. Is it possible that the speechreading instrument in the present study was not sensitive enough to differentiate the two groups as far as accuracy measure is concerned? Future studies need to assess speechreading performance at different psycholinguistic levels in Chinese, as it has been done for English (Mohammed *et al*., 2006; Tye-Murray *et al*., 2014) and Brazilian Portuguese (Oliveira *et al*., 2014). Mohammed *et al*. (2006), for example, not only found an advantage of the deaf in speechreading, but also found that deaf and hearing adults speechread single words more accurately than sentences, which were easier than short stories. These results led Mohammed *et al*. (2006:628) to suggest that 'the use of appropriate task parameters, including selected vocabulary and syntax and appropriate response mode, allows perceptual speechreading differences between individuals to emerge clearly and that then, the effects of hearing status becomes evident'.

An alternative explanation for the lack of speechreading differences between hearing and hearing impaired sixteen year olds in China might be related to their decreased exposure to the oral channel. It has commonly been observed that long experience with visual speech improves visual speech perception - even when audiovisual experience is minimal, as is the case for individuals who are deaf and hard of hearing. Auer and Bernstein (2008), in a study estimating when and how words are acquired by deaf individuals, observed that their deaf participants who subjectively reported relying more on the spoken channel for acquisition also performed better on an objective measure of visual speech perception. However, the hearing impaired participants in the present study had just started high school at the time of data collection, and they were required to spend more time studying the textbooks of different subjects alone in order to prepare for the competitive college entrance examination. This may have contributed to the decline in the speechreading performance. To some extent, the HI participants might have greater need for visual language communication and have consequently developed better exploitation of visual speech information at a younger age in comparison to their hearing peers; however, they began to lose this advantage as the increasing demand for reading written material began at the start of high school. To explore this possibility, we are currently comparing the development of speechreading performance in Chinese-speaking children and adults with HL to those with NH.

Another important finding of the present study involved both deaf and hearing individuals. The characteristics of sounds to be identified, influenced speechreading performance. Specifically, initials (consonants) in Chinese were identified less accurately and more slowly than simplex finals and complex finals. This finding is consistent with a previous observation that people with high-frequency sensorineural hearing loss typically demonstrate consonant confusions but have less difficulty with vowel identification (Lesner, Sandridge, & Kricos, 1987). This vowel advantage may reveal the significance of the vowel/consonant contrast for visual speech memory (De Gelder & Vroomen, 1994), or it may reflect a greater degree of phonological ambiguity inherent in visually perceiving (initial) consonants in comparison to the two types of vowel sounds (Chen & Lei, 2017).

In conclusion, the present study contributes to our understanding of speechreading ability in school-aged deaf and typically hearing children. In particular, the results of the present study show that speechreading skills in Chinese speakers are influenced by hearing status, characteristics of sounds to be identified, as well as the speechreading measures used.

Acknowledgements

This work was supported partially by funding from the China National Social Science Foundation (grant number 15BYY069). We would like to thank the participants, their families, and the teachers involved in conducting this research.

References

Arnold, P., & Köpsel, A. (1996). Lipreading, reading and memory of hearing and hearing-impaired children. *Scandinavian Audiology*, *25*, 13-20. doi:10.3109/01050399609047550

Auer, E. T., Jr., & Bernstein, L. E. (2007). Enhanced visual speech perception in individuals with early-onset hearing impairment. *Journal of Speech, Language, and Hearing Research*, *50*, 1157-1165. doi:50/5/1157 [pii] 10.1044/1092-4388(2007/080)

Auer, E. T., Jr., & Bernstein, L. E. (2008). Estimating when and how words are acquired: A natural experiment on the development of the mental lexicon. *Journal of Speech, Language, and Hearing Research*, *51*(3), 750-758.

Bernstein, L. E., Demorest, M. E., & Tucker, P. E. (2000). Speech perception without hearing. *Perception and Psychophysics*, *62*, 233-252. doi:10.3758/BF03205546

Conrad, R. (1977). Lip-reading by deaf and hearing children. *British Journal of Educational Psychology*, *47*, 60-65. doi:10.1111/j.2044-8279.1977.tb03001.x

Chen, L., & Lei, J. (2017). The development of visual speech perception in Mandarin Chinese-speaking children. *Clinical Linguistics and Phonetics*, *31*(7-9), 514-525. doi: 10.1080/02699206.2016.1277391

De Gelder, B., & Vroomen, J. (1994). Memory for consonants versus vowels in heard and lipread speech. *Journal of Memory and Language*, *31*, 737-756.

Elphick, R. (1996). Issues in comparing the speechreading abilities of hearing-impaired and hearing 15 to 16 year-old pupils. *British Journal of Educational Psychology*, *66*, 357-365.

Gagné, J. -P., Charest, M., Le Monday, K., & Desbiens, C. (2006). Evaluation of an audiovisual-FM system: Speechreading performance as a function of distance. *International Journal of Audiology*, *45*(5), 295-300.

Green, W. B., Green, K. W. & Holmes, D.W. (1981). Growth of speechreading proficiency in young hearing-impaired children. *The Volta Review*, 83, 389-393.

Kyle, F. E., Campbell, R., Mohammed, T., Coleman, M., & MacSweeney, M. (2013). Speechreading development in deaf and hearing children: Introducing the test of child speechreading. *Journal of Speech, Language, and Hearing Research*, *56*(2), 416-426.

Kyle, F. E., Campbell, R., & MacSweeney, M. (2016). The relative contributions of speechreading and vocabulary to deaf and hearing children’s reading ability. *Research in Developmental Disabilities*, *48*, 13-24.

Lei, J., & Chen, L. (2015). Speech-reading performance of hearing impaired students in China. *Speech, Language, Hearing*, *18*(4), 188-195.

Lesner, S., Sandridge, S., & Kricos, P. (1987). Training influences on visual consonant and sentence recognition. *Ear and Hearing*, *8*(5), 283-287.

Lyxell, B., & Holmberg, I. (2000). Visual speechreading and cognitive performance in hearing-impaired and normal hearing children (11-14 years). *British Journal of Educational Psychology*, *70*, 505-518. doi:10.1348/000709900158272

Mohammed, T., Campbell, R., MacSweeney, M., Barry, F., & Coleman, M. (2006) Speechreading and its association with reading among deaf, hearing and dyslexic individuals. *Clinical Linguistics and Phonetics*, *20*(7-8), 621-630, DOI: 10.1080/02699200500266745

Myklebust, H. R., & Neyhus, A. I. (1970). *Diagnostic test of speechreading*. New York: Grune and Stratton.

Oliveira, L. N., Soares, A. D., Chiari, B. M. (2004). Speechreading as a communication mediator. *CoDAS*, *26*(1), 53-60.

Tye-Murray, N., Hale, S., Spehar, B., Myerson, J., & Sommers, M. (2014). Lipreading in school-age children: The roles of age, hearing status, and cognitive ability. *Journal of Speech, Language, and Hearing Research*, *57*, 556-565.

To cite this article: Chen, L., Lei, J., & Gong, H. (2017). Speechreading performance of Chinese adolescents with and without hearing impairment. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 85-91). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Mixed relative clauses: Some consequences for code-switching theories

Gloria Cocchi, Cristina Pierantozzi
gloria.cocchi@uniurb.it, cristina.pierantozzi@uniurb.it

University of Urbino

Abstract. This work aims at testing the possibility of code-switching between the DP-head of a relative clause and the relative pronoun in mixed German-Italian sentences. In particular, we will see if differences in acceptability emerge when the main clause is in Italian rather than in German, and with different mixed agreement patterns. Data have been obtained through an Acceptability Judgment Task administered to some 2L1 adults and teenagers, in order to see if different responses are given by the two different age groups. Finally we will discuss the implications that the results obtained bring about for code-switching theories.

Keywords: code-switching, relative clauses, selected gender, analogical gender

Introduction (1)

Changing language along a conversation is a natural and rule-governed phenomenon typical of bilingual speakers. This linguistic strategy, termed code-switching (CS), is now at the centre of a lively and intense field of research covering aspects related to the two basic dimensions, E-Language (performance) and I-Language (competence). One of the most debated points is whether CS should be distinguished from borrowing and, if so, which type of process speaks in favour of a borrowing analysis (i.e. morpho-syntactic and/or phonological integration).

In this paper we aim at verifying how fine-grained the morphological integration in CS sentences may be by manipulating the language of functional heads in German-Italian mixed relative clauses. We will use the formal tools of generative grammar and assume, in line with the Null Hypothesis on Code-Switching, that 'code-switching and pure languages are governed by the same set of constraints and principles in syntax, production and pragmatics' (Chan, 2003:17). We will also discuss the theoretical implications brought about by our data for current models of bilingual competence.

This work proceeds as follows. In the next section, we describe the various types of German-Italian mixed relative clauses and the problems that arise concerning mixed agreement patterns. The following section presents theoretical problems raised by these data and predictions that current theories on code-switching would make. Subsequently, we outline the Acceptability Judgment Task (AJT), as well as some examples of test sentences, followed by a section analysing the results obtained. In the final part, we discuss some theoretical implications for CS theories.

Mixed relative clauses (2)

Relative clauses (RCs) are complex sentences which involve agreement outside DP, i.e. long-distance agreement between the external DP, namely the head of the RC and the relative pronoun (RP).

Several interesting issues arise. To start with, RCs are complex sentences involving syntactic operations which have a high cost for processing. We may thus wonder whether the complexity of RCs, in terms of processing cost, has a direct impact on the availability of a code-switching point in a mixed sentence. Furthermore, the N-head of the RC establishes a syntactic relation with two functional categories, the external D and the relative pronoun/operator. This 'double' syntactic life of the N-head is even more evident in languages like German, where the external D and the RP – also a D – agree in gender and number but may carry a different Case feature:

(1) Ich kenne den Man, der uns heute besucht hat
I know the(m.sg. ACC) man, that(m.sg. NOM) us today visited has
'I know the man who has visited us today'

The syntactic relation of the N-head with both the external D and the RP raises interesting questions for CS theory, especially when the head of the RC is a mixed DP. In this specific case, we might select from Italian either the N-head, as in (2.a), or the external D, as in (2.b):

(2) a. Ich kenne den uomo, der uns heute besucht hat
b. Ich kenne il Man, der uns heute besucht hat

Furthermore, the main clause may be either in German, as in (2.a-b), or in Italian, as in (3.a-b):

(3) a. (Io) conosco den uomo, der uns heute besucht hat
b. (Io) conosco il Man, der uns heute besucht hat

Finally, when a noun has different genders in the two languages (e.g., German *sessel* (M) vs. Italian *poltrona* (F) 'armchair'), we have four possible agreement patterns:

SGD (selected gender of the determiner): gender match between D, N and RP (4a);
SGN (selected gender of the noun): gender match between D and N, mismatch with RP (4b);
AGD (analogical gender of the determiner): gender match between D and RP, mismatch with N (4c);
AGN (analogical gender of the noun): gender match between N and RP, mismatch with D (4d):

(4) a. Il sessel, der .... / Die poltrona, die….
b. Il sessel, die ..... / Die poltrona, der….
c. La sessel, die .... / Der poltrona, der....
d. La sessel, der .... / Der poltrona, die....

The various mixed RCs discussed so far display different degrees of morpho-syntactic integration and thus prove particularly suitable for disentangling the central question of borrowing vs. CS.

Theoretical implications (3)

The study of CS grammar has been guided by the basic assumption that switching points are available where the two languages of the bilingual share feature specifications of the words involved (i.e. categorical features, word order and phi-features) or where no conflict arises (i.e. adjuncts, sentence boundaries). Although this underlying idea has found robust empirical support (cf. Alternation, in Muysken 2000), published data clearly show that feature compatibility is not a strict and systematic requirement, and the conflict between the two grammars is often resolved in favour of the language which provides the morpho-syntactic structure of the sentence. As we will see below, the grammatical constraints reported in the literature make different predictions about the grammaticality of mixed RCs.

Performance accounts

According to the Matrix Language Frame Model (MLF; Myers-Scotton 1993 and subsequent work), all functional categories in CS should be provided by only one of the two languages in contact (the matrix language), while the other language (embedded language) can only provide lexical material. The mixed Italian/German RCs seen in examples (2-3) above display different degrees of morpho-syntactic integration. In particular, only the mixed RC in (2.a) is perfectly accounted for under this theory, since all functional categories come from one language (German). On the other hand, the other types of sentences represent violations of the MLF principles, as functional categories are expressed in different languages: T and external D in Italian, RP in German in (3.b); T in Italian, external D and RP in German in (3.a); external D in Italian, T and RP in German in (2.b). Regarding the cases in (4), since the RP is always in German, only mixed DPs with German D and Italian N are compatible with MLF (irrespectively of the agreement pattern), provided the main clause is in German too (2.a).

The Borrowing Hypothesis (BH; Poplack & Meechan, 1995) assumes instead that all occurrences of word insertion within a sentence in a different language are to be analysed as borrowings, as far as the words are perfectly integrated under a morpho-syntactic point of view. Therefore, under this theory some of the mixed agreement patterns in (4) would not be acceptable, in particular SGN (4b), where the RP agrees only with the equivalent noun, and AGN (4d), where there is a gender mismatch between D and RP. Crucially, in the latter case, the external noun controls an agreement relation outside DP (with RP) and, hence, it cannot be treated as a pure borrowing.

Generative approaches

In generative approaches, a mixed sentence is predicted to be well-formed if the feature checking process successfully deletes uninterpretable features. There are no special constraints, provided that the requirements of the mixed grammar are respected. Therefore, only SGD is predicted to be grammatical, while all other combinations are not.

However, the AGD type has been widely attested in the speech of both bilingual children and adults (Liceras, Fernández, Perales, Pérez-Tattam, & Spradlin, 2008). Again, the question to be answered is whether this problematic datum should be considered borrowing or not. Two lines of research are discussing the problem: the Bi-lexical Model of bilingual competence (MacSwan 1999) and the Language Synthesis Model (LSM; Lillo-Martin, Müller de Quadros, & Chen Pichler, 2016). The former assumes a lexicalist view of the lexicon and translates BH into minimalist terms (in line with Chomsky, 1995). The latter rejects the lexicalist hypothesis and, along with Distributed Morphology (DM; Halle & Marantz, 1993), assumes instead that lexical items are void of phonological content, which is inserted into the terminal nodes only later, after syntax.

Thus, the various mixed agreement types, which are by-products of morphology-phonology mapping, are in principle all derivable. Restrictions are dictated by the Subset Principle (Halle, 1997), which licenses the insertion of the phonological content of functional heads (external D and RP). Insertion takes place only if: a) the feature specification of the phonological content matches the feature specification of the abstract structure (5), or b) the phonological content is underspecified with respect to the feature specification (6):

(5) $[_{DP}$ D phi: SG, F, $[_{NUM}$: SG $[_{n\ GEN}$: F [√matita]]]] $[_{CP}$ D phi: SG, F...
$die_{sg.f}$ matita, $die_{sg.f}$ (SGD)
$la_{sg.f}$ Bleistift, $die_{sg.f}$ (AGD)

(6) $[_{DP}$ D phi: SG, F, $[_{NUM}$: SG $[_{n\ GEN}$: F [√casa]]]] $[_{CP}$ D phi: SG, F...
$la_{sg.f}$ Haus, $das_{sg.n}$ (AGN)
$die_{sg.f}$ casa, $das_{sg.n}$ (SGN)

Research questions

In order to see if the predictions made by the various theories are confirmed, we will address the following Research questions:

1. Are all of the mixed agreement types in (4) equally accepted by 2L1 speakers, and at which percentage?
2. Do differences in acceptability emerge when the language of the main clause is Italian (see (3)) or German (see (2))? In particular, does bilingual competence tolerate the violations discussed in 3.1 above?
3. Does the age of 2L1 speakers play any role?

The data (4)

In order to provide answers to these research questions, we administered online an acceptability judgment task (AJT) to some Italian-German bilingual speakers. We selected nine 2L1 participants: three adults and six teenagers (average age 12), the latter attending the Europäische Schule München, a bilingual Secondary School. The AJT (3 Scale: acceptable, marginally acceptable and unacceptable, with corrections) contains test sentences that were provided in random order and divided into different forms, each of which includes 30% of distracters.

The questionnaire (see Table 1) has been specially designed to test different issues at the same time: a) the participants' competence in gender agreement; b) the acceptability of the switching point between the DP-head and the RC; c) whether the language of the main clause plays any role; d) if all types of mixed agreement are accepted, and at which percentage.

Table 1. Test sentences: Overview of gender distribution

Gender of the N pair	*m/m*	*f/f*	*m/f*	*f/m*	*n/f*	*n/m*	*TOT*
Monolingual German RC							
MATCH			2	2	2	2	8
GENDER TRANSFER			2	2	2	2	8
Mixed RC							
MISMATCH	6	6					12
MATCH	6	6					12
Type of mixed agreement							
SGD			2	2	2	1	7
AGD			2	2	2	2	8
AGN			2	2	2	2	8
SGN			2	2	2	1	7
TOT	12	12	12	12	12	10	

We provide below some examples of test sentences. Each group corresponds to one of the different mixed agreement patterns seen in (4). RC is obviously always expressed in German. Within each group, example (a) features a German main clause where an Italian N-head is inserted (as in (2a)); in example (b) the D-head and the main clause are in Italian (cf. (3b)); in example (c) the N-head and the main clause are in Italian (cf. (3a)). We have not yet tested mixed sentences like (2b) above, where only an Italian D is inserted into an otherwise monolingual German sentence.

(7) **SGD (gender match between D, N and RP)**

a. Die gonna, die Maria heute anzieht, ist gelb
'the(F) skirt(F) that(F) Maria today wears, is yellow'

b. Il Licht, der im Himmel blitzt, è solo un aereo
'the(M) light(M), that(M) in-the sky flashes, is only an airplane'

c. Ho buttato den topo, den der Hund versteckt hatte
'I-have thrown-away the (M) mouse(M), that(M) the dog hidden had'

(8) **AGD (gender match between D and RP)**

a. Ich will die racconto lesen, die von unserem Abenteuer erzählt
'I will the(F) story(M) read, that(F) of your adventure tells'

b. La Rock, die Maria heute anzieht, era di tua madre
'the(F) skirt(M) that(F) Maria today wears, was of your mother'

c. Voglio mettere das camicia, das Martin anprobieren möchte
'I-want to-put-on the(N) shirt(F), that(N) Martin try-on wanted'

(9) **AGN (gender match between N and RP)**

a. Woher kommt das acqua, die auf dem Boden ist?
'where comes the(N) water(F) die(F) on the floor is'

b. Mi prendi il Buch, das Johann online bestellt hat?
'to-me bring the(M) book(N) that(N) Johann online ordered has'

c. Ecco die muro, den dein Freund gemalt hat
'here-is the(F) wall(M) that(M) your friend painted has'

(10) **SGN (gender match between D and N)**

a. Das ist die stella, der uns den Weg zeigen wird
‘this is the(F) star(F) that(M) to-us the way show will’
b. Ecco il Sessel, die uns heute geliefert wurde
‘here-is the(M) armchair(M) that(F) today delivered was’
c. Mi passi den pane, das Maria mitgebracht hat?
‘to-me pass the(M) bread(M) that(N) Maria brought has’

Results (5)

The figures below show the responses provided by the bilinguals.

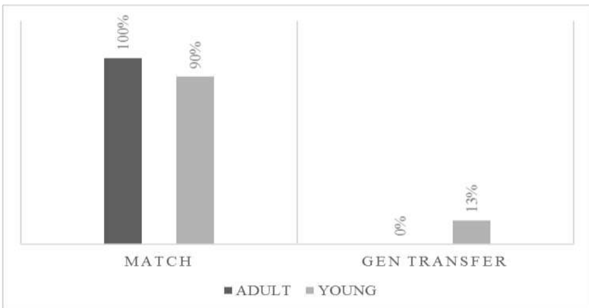


Figure 1. Competence in gender agreement in monolingual German RCs

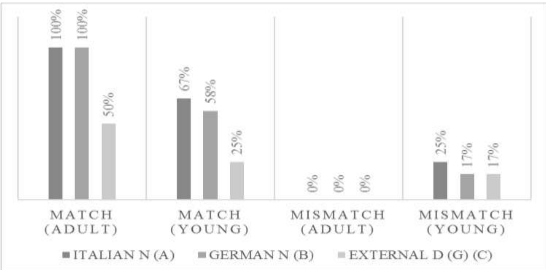


Figure 2. Switching point across mixed RC types

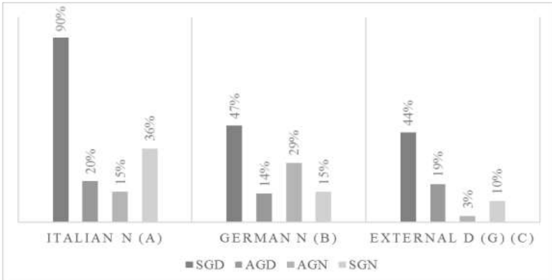


Figure 3. The distribution of mixed agreement patterns across mixed RCs

In Figure 1, we see that all bilinguals show high competence in monolingual gender agreement. Interestingly, gender transfer (exemplified in (11) next) is completely rejected by adults but marginally accepted by teenagers, though the percentage is rather low:

(11) Eine Wochezeitung ist eine Zeitung, der jede Woche einmal erscheint,
‘A weekly-newspaper is a(F) newspaper(F) that(M) comes out once a week ‘
German N: *zeitung*(F); Italian equivalent N: *giornale*(M)

From Figure 2, we infer that the availability of the switching point between the DP-head and the RC changes on the basis of the language of the main clause: it is higher when the language of the main clause matches the language of the external D. Besides, young 2L1 competence in mixed agreement is poorer: teenagers show a lower rate of acceptance of match sentences and, at the same time, they sometimes do not reject gender mismatch, at least with a neuter RP (indeed neuter can be considered as the default gender); see (12):

(12) die magia, das mich verzaubern kann, non esiste!
‘the(F) magic(F), that(N) might enchant me, does not exist!’
Italian N: *magia*(F); German equivalent N: *zauberei*(F)

From the data in Figure 3, we can conclude that the SGD pattern is overall the preferred one, and it shows a very high rate of acceptance in the (a) type (insertion of Italian N) for all speakers. Interestingly, the acceptance rates of SGN are relatively high in the (a) type, without gender distinctions, but only for young 2L1 (58%), while adults definitely reject this pattern. The rates of acceptance of the AGD pattern is much lower than the SGD one, even in the RC type, where it represents the highest degree of morphological integration (insertion of Italian N). Finally, the rates of acceptance of the AGN mixed agreement pattern are higher when the N-head is expressed in German rather than in Italian, especially for adults.

We can now provide an answer to our Research questions (cf. 3.3. above).

Research question 1:

The switching point between the DP and the RC is widely accepted (see Figure 2 on match and mismatch). However, the percentages of acceptability of the four mixed agreement types varies dramatically (Figure 3), with the SGD type strongly preferred with respect to all the others, as expected.

Research question 2:

The language of the main clause plays an important role. Indeed, all speakers tend to prefer the cases in (a), where the Italian N-head is inserted into an otherwise monolingual German sentence. Moreover, the language of the main clause also influences the acceptance rates of mixed agreement patterns: when the main clause is in German, SGN is preferred with respect to AGD and AGN. When the main clause is in Italian, SGN is almost always rejected, while the AGN acceptance rate increases.

Research question 3:

We noted some interesting discrepancies in the responses given by adults and teenagers:

- teenagers' competence in agreement is somehow lower: unlike adults, they sometimes accept gender transfer (Figure 1) and gender mismatch in mixed agreement (Figure 2);
- teenagers show a non-negligible rate of acceptance of SGN, especially with the insertion of an Italian N, while this pattern is almost always rejected by adults (see also Cocchi & Pierantozzi, 2016);
- as for AGN, while adults accept it only with a German N, teenagers accept it also with an Italian N.

Conclusion: Implications for CS theories (6)

MLF is only partially confirmed by our data: match sentences are most widely accepted, which is expected. However, the acceptance rate does not seem to vary when functional categories are not all expressed in the same language (German), and in particular when the external D is in Italian (3.b). Besides, under this theory the AGD mixed agreement pattern should perform much better than it does, at least in sentences like (8a), where all functional categories belong to the same language (German).

The non-negligible rate of acceptance of the AGN type (where the RP agrees with the N-head of the RC, but not with the external D) proves instead problematic for BH, as this theory does not account for the possibility of a long-distance agreement outside DP (see Cocchi & Pierantozzi, 2016, 2017). Likewise, the Bi-lexicalist model to CS predicts the high rate of acceptance of the SGD type, since the functional D-heads (external D and RP) are checked and valued by the selected N. However, under this approach, the AGD, AGN and SGN patterns would all be problematic.

Hence, our data seem to support a DM approach, as depicted in LSM, which assumes that the various mixed agreement patterns are in principle all derivable. However, this approach raises the non-trivial problem of the overlapping output generable by the bilingual I-language (Pierantozzi, 2016). The data we collected seem to suggest thus far that this approach is on the right track, provided that:

- before syntax, the abstract morphemes are selected basically from one language. In fact, although not systematically, the mixed RCs in the (a) examples are preferred with respect to the other cases;

- in mixed agreement, gender restrictions apply according to the underspecification condition. Indeed we recorded a higher rate of acceptability of mixed agreement types (especially AGN) involving the default gender (neuter);
- external factors, such as the age of bilingual speakers, may influence the control of the two languages before and after syntax.

Declaration of authorship

This work is the result of the collaboration of the two authors in all respects. Nevertheless, Gloria Cocchi contributed to Sections 2, 5 and 6, and Cristina Pierantozzi contributed to Sections 1, 3 and 4.

References

Chan, B. (2003). *Aspects of the syntax, the pragmatics, and the production of code-switching: Cantonese and English* [Berkeley Insights in Linguistics and Semiotics 51]. New York: Lang.

Chomsky, N. (1995). *The minimalist program*. Cambridge MA: MIT Press.

Cocchi, G., & Pierantozzi, C. (2016). Code-switching in relative clauses: Some theoretical implications. *Quaderni di Linguistica e Studi Orientali*, *2*, 95-119.

Cocchi, G., & Pierantozzi, C. (2017). Nominal gender and agreement relation outside the DP domain in code-switching contexts. *Lingue e Linguaggio*, *1*, 35-62.

Halle, M. (1997). Distributed morphology: Impoverishment and fission. In B. Bruening, Y. Kang & M. McGinnis (eds.), *MIT Working Papers in Linguistics 30: Papers at the Interface* (pp. 425-449). Cambridge, MA: MITWPL.

Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 111-176). Cambridge, MA: The MIT Press.

Liceras, J. M., Fernández, R., Perales, S., Pérez-Tattam, R., & Spradlin, K. T. (2008). Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional-lexical mixings. *Lingua*, *118*, 827-851.

Lillo-Martin, D., Müller de Quadros, R., & Chen Pichler, D. (2016). The development of bimodal bilingualism: Implications for linguistic theory. *Linguistic Approaches to Bilingualism*, *6*(6), 1-63.

MacSwan, J. (1999). *A Minimalist approach to intrasentential code switching*. New York: Garland.

Muysken, P. (2000). *Bilingual speech*. Cambridge, MA: Cambridge University Press.

Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in codeswitching*. Oxford, UK: Oxford University Press.

Pierantozzi, C. (2016). The language synthesis model and the problem of the invisible derivation. *Linguistic Approaches to Bilingualism*, *6*(6), 804-807.

Poplack, S., & Meechan, M. (1995). Orphan categories in bilingual discourse: Adjectivization strategies in Wolof-French and Fongbe-French. *Language Variation and Change*, *7*(2), 169-194.

To cite this article: Cocchi, G.., & Pierantozzi, C. (2017). Mixed relative clauses: Some consequences for code-switching theories. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 92-98). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Second dialect and second language imitation of geminates by Colombian Spanish speakers

Daniel Varona Cordero, Esperanza Ruiz-Peña, Elkin Dario Sierra,
Ryan Andrew Stevenson, Yasaman Rafat
dvaronac@uwo.ca; mruizpen@uwo.ca; esierrar@uwo.ca; rsteve28@uwo.ca; yrafat@uwo.ca

Western University

Abstract. The present study compares second language (L2) and second dialect speech (D2) speech learning by examining geminate imitation of Standard Italian (L2) and Havana Cuban Spanish (D2) by native Colombian Spanish-speaking participants. We test the effect of both language and salience (length). We provide new data that show that participants were able to produce geminates from a phonological point of view, though they had difficulty producing them in a target-like manner in both the L2 and D2 contexts, when we compared geminate duration. Moreover, length was the main predictor of non-target-like productions, where shorter geminates were easier to imitate. It is proposed that there might be a trade-off between acoustic salience and target-like productions in L2 and D2 speech learning.

Keywords: second language/dialect imitation, phonetics, salience, production, Spanish, Italian

Introduction

Whether proximity/similarity makes a second dialect (D2) more or less acquirable or not has not been resolved yet (e.g., Siegel, 2010). Not much is known about whether it is easier to learn new sounds in a second language (L2) or in a D2 either, despite much interest in D2 speech learning (e.g., Babel, 2010; Nielsen, 2011). The only study to our knowledge that has compared L2 and D2 imitation has been Ruiz-Peña, Sevilla, and Rafat (2018). With the aim of determining whether L2 and D2 speech learning operate in the same way, the D2 production of Ecuadorian Spanish assibilated rhotics by native Sevillian Spanish speakers in a real and nonce word imitation tasks were compared with the findings on the production of Mexican assibilated rhotics by native English-speaking participants in auditory only and auditory-orthographic conditions, previously reported in Rafat (2015). Similarly to Rafat (2015), only a small percentage of assibilated rhotics were actually produced as assibilated rhotics by Andalusian Spanish-speaking participants when imitating Ecuadorian Spanish (Ruiz-Peña *et al*., 2018). Moreover, assibilated rhotics in the real word imitation task in the D2 context were mainly realized as rhotics in the D2 context, as in the auditory-orthographic condition in the L2 study by Rafat (2015). However, in the nonce word imitation task in the D2 context, assibilated rhotics were mainly produced as sibilants by the native L2 learners in the auditory-only group in Rafat (2015). It was concluded that at the onset of acquisition, D2 and L2 acquisition for the most part appear to operate in a similar manner, where assibilated rhotics may be categorized as a 'similar' sound in Flege's (1995) terms and may be mapped on to their nearest L1 category due to equivalence classification.

A length contrast between singleton (short) and geminate (long) consonants is observed in Standard Italian (e.g., /sete/ 'thirst' vs. /setːe/ 'seven') (Payne, 2005). Gemination also exists in Cuban Spanish varieties, although it has been described as allophonic only and as a result of an assimilatory process (Carlson, 2011). This said, the assimilation process can result in minimal pairs such as /paɾto/ [patːo] 'birth' and /pato/ [pato] 'duck' in this variety of Spanish Moreover, gemination is not generally considered to be a phonological feature of Spanish. This is despite the fact that a length contrast exists between the tap and the trill (e.g., *perro* /pero/ 'dog' vs. *pero* /peɾo/ 'but') in all varieties of Spanish, even when trilling is absent (see Colantoni & Rafat, 2013). Geminate acquisition has been studied from various perspectives such as child acquisition (Khattab & Al-Tamimi, 2015), L2 speech learning (Motohashi-Saigo & Hardison, 2009; Sorianello, 2014), and in heritage and attrited speakers (Celata

& Cancila, 2010; Rafat, Mohaghegh, & Stevenson, 2017). With respect to L2 speech learning, previous studies provide evidence for non-native like rate of geminate production (Sorianello, 2014) and/or non-target-like geminate productions (Mah & Archibald, 2003; Sorianello, 2014).

The current study has two aims. First, it will investigate whether imitating geminates (long sounds) is easier in an L2 or a D2 context. Specifically, it will examine the imitation of Standard Italian and Havana Cuban Spanish (here on referred to as Cuban Spanish) by native Colombian Spanish-speaking participants. Second, it will examine the role of salience (length) in geminate imitation in both contexts. That salience plays a role in L2 speech learning is well known (Colantoni & Steele, 2008; Flege, 1995; Rafat, 2011; 2015). Flege's Speech Learning Model postulates that a larger acoustic-phonetic distance between the first language (L1) and L2 sounds would result in the establishment of a new category. The absence of a large acoustic-phonetic distance between L1 and the L2 sounds would lead to equivalence classification, where the L2 sound would be mapped on to its nearest L1 category. The role of salience has also been examined in D2 studies that have focused on phonetic accommodation (Babel, 2010; MacLeod, 2012) and there is generally consensus that length is an inherently salient acoustic property of sounds and may also be salient to Spanish-speaking learners (e.g., Escudero, 2001).

Research questions and predictions

The research questions in this study are as follows: (1) Will participants be able to imitate geminates in the L2 and D2 conditions? (2) Does language modulate the degree of target-like productions? Specifically, will it be easier to learn geminates in an L2 (Standard Italian) or a D2 (Cuban Spanish) context? And (3) Does salience (length) affect D2 and L2 geminate imitation?

It is predicted that geminate imitation will pose difficulty (e.g., Giannini, 2003; Mah & Archibald, 2003; Sorianello, 2014) for the participants in both the L2 and D2 conditions (H1). Based on Ruiz-Peña *et al*., (2018), it is also predicted that geminate imitation will be equally difficult in both the L2 and D2 conditions, at the onset of acquisition in native participants (H2). Because length is salient (e.g., Escudero, 2001) longer geminates will be more salient than shorter ones, and thus easier to imitate in both the L2 and D2 conditions.

Method

Participants

Participants were 3 male and 7 female Colombian Spanish-speakers, who were all educated and lived in Ontario, Canada. Their ages ranged between 21 and 43 years old (mean age = 35). None of the participants had previous contact with the Italian language, and they had very little contact with the Cuban variety of Spanish.

Procedures

Participants performed an imitation task in both the L2 and D2 and completed a language background and a language attitude questionnaire. These tasks were completed within a single experiment session, which lasted two hours and thirty minutes. Participants received a brief training session that lasted 15 minutes prior to the start of the main task. Training consisted of telling the participant to listen to the carrier sentences, pay attention to the target words carefully and imitate the target words as closely as possible while repeating the entire carrier phrases. The carrier phrase in Italian was *Dico* ______*un' altra volta* 'I say ________once again' and in Spanish *Digo* ________ *una y otra vez* 'I say ________ once and again'. We used Praat to manipulate the duration of geminates of the 90 target stimuli (both the Italian and Cuban Spanish) and created 6 different conditions (see Table 1).

Both the training and the main task were done via PowerPoint. The instructions were also the same in both the training and the main task. In each condition, 45 words were presented individually using the

Spanish and Italian carrier phrases. Participants were auditorily presented with each phrase three times in a row, before they had to imitate them. The inter-stimuli interval was 3 seconds and participants were told to imitate them immediately after the third presentation of each phrase.

Table 1. Description of the six conditions in the experiment: Short L2, unaltered L2, long L2, short D2, unaltered D2 and long D2

Condition	Geminate length manipulations
Short L2 (Standard Italian)	Minimum length of Cuban Spanish tokens
Unaltered L2 (Standard Italian)	Natural length in Standard Italian
Long L2 (Standard Italian) Long	Maximum length of Cuban Spanish tokens
Short D2 (Cuban Spanish)	Minimum length of Standard Italian
Unaltered Cuban Spanish)	Natural length in Cuban Spanish
Long D2 (Cuban Spanish)	Maximum length of Standard Italian tokens

Stimuli

The stimuli consisted of 45 Spanish real words, and 45 Italian real words that were first produced in Standard Italian and Cuban Spanish and were then manipulated for length to create short, and long geminates in addition to unaltered geminates. The stimuli contained 9 different geminate sounds including voiced and voiceless stops ([bː dː gː] and [pː tː kː], respectively), and sonorants ([mː lː nː]) in both Cuban Spanish and Standard Italian. All target words were controlled for position and stress, number of syllables, and the previous vowel. Geminates were included inter-vocalically in an unstressed position preceded by the vowel [a] in disyllabic words (e.g., Cuban Spanish ['abːa] (/alba/ in Spanish), Standard Italian ['sabːa]).

Data analysis and results

Analysis procedure

8100 tokens were produced by the participants. Consonant duration in these tokens together with those used for the recordings (270 Standard Italian words and 270 Havana Spanish Cuban words) were analyzed both auditorily and acoustically using Praat and were compared.

Mean geminate duration of participants' productions were calculated per token, per condition. To assess error magnitude between duration of the stimulus and the participants' actual production for each geminate, the Euclidean distance between the 5 tokens (each consonant had 5 tokens) was calculated in SPSS. This resulted in a single, non-directional dissimilarity variable for each consonant, in each of the 6 listening conditions. Next, a 2x3x9, mixed-model ANOVA was conducted with Language (Cuban Spanish (D2) or Italian (L2)), duration condition (short, unaltered, long), and consonant ([b p d t g k l n m]) as factors. A Greenhouse-Geisser correction was applied to all comparisons to account for violations of assumptions of sphericity. Follow-up hierarchical regressions were run including all independent variables associated with significant main effects.

Analysis of variance results

Figure 1 shows that geminates were not produced in a target-like manner. For the most part, there was evidence of undershoot, but there was also some evidence of overshoot/hyper-articulation in the 'short' and 'unaltered' conditions. The three-way ANOVA with factors of language, duration condition, and consonant was conducted to measure the impact of each on how well a participant was able to replicate geminate duration (Figure 5). A main effect of language was found, such that errors were greater with L2 than with D2 ($F(1,9) = 1{,}707.22$, $p < 0.001$, $\eta p2 > 0.99$). Second, a main effect of duration condition was observed, such that greater errors were observed with longer geminate durations ($F(2,18) = 176.94$, $p < 0.001$, $\eta p2 = 0.95$). Third, a main effect of consonant was observed, with different consonants being associated with varying degrees of error magnitude ($F(8,72) = 60.50$, $p < 0.001$, $\eta p2 = 0.87$). All two-way interactions within the three-way ANOVA were also significant.

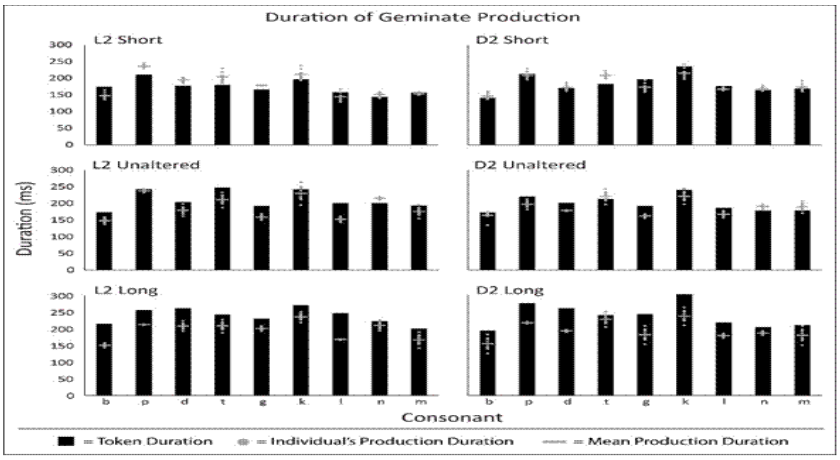


Figure 1. Durations of geminate production. Mean token durations for each geminate produced by the native speaker of Italian and Spanish are depicted as black columns. Mean duration of participants' productions are depicted as grey lines, with each individual's mean duration depicted as a grey circle. L2 refers to Italian, a second language, and D2 refers to Cuban Spanish, a second dialect

A language-by-duration condition interaction was observed, such that the greatest differences in error magnitude between languages was observed in the unaltered condition, and with differences in error magnitude being smallest with longer durations (Figure 2; $F(2,18) = 17.44$, $p = 0.002$, $\eta p2 = 0.66$). Likewise, significant language-by-consonant ($F(8,72) = 48.13$, $p < 0.000$, $\eta p2 = 0.84$) and duration-by-consonant ($F(16,144) = 32.07$, $p < 0.001$, $\eta p2 = 0.78$) interactions were observed. Effects of language and duration condition for each consonant individually can be seen in Figure 3.

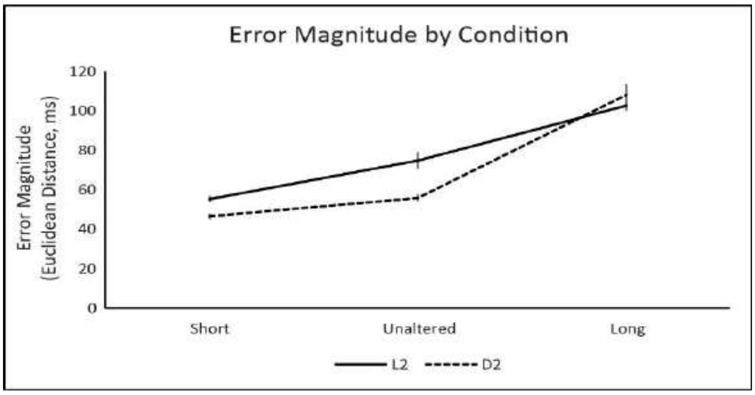


Figure 2. Mean error rate across conditions. Magnitude of duration errors increased in longer-duration conditions, across languages and averages across consonants. L2 refers to Italian, a second language, and D2 refers to Cuban Spanish, a second dialect. Error bars represent standard error of the mean

.

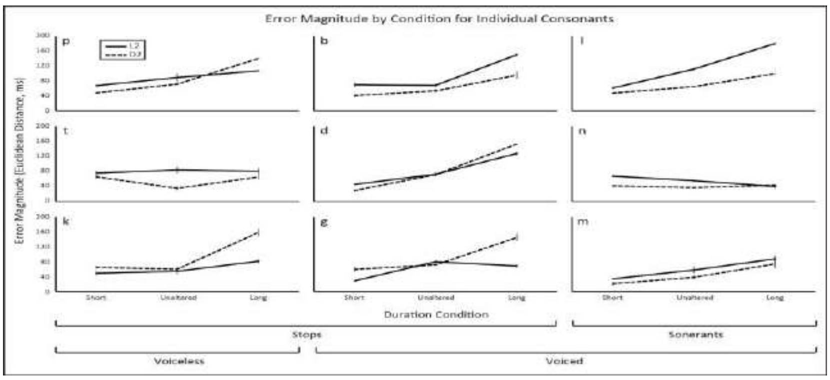


Figure 3. Error rates of individual consonants across conditions. L2 refers to Italian, a second language, and D2 refers to Cuban Spanish, a second dialect. Error bars represent standard error of the mean

Hierarchical regression results

To explore whether our three variables of interest accounted for distinct portions of the variance within the dependent measure, we conducted a 2-step hierarchical regression. The fist model included individual participants as predictors to control for any between-subject differences. The second model added language, duration condition, and consonant into the model to test for experimental effects. The first model, which controlled for individual differences, was not significant, though it trended towards significance ($p = 0.062$). The second model was significant overall ($p < 0.001$), and revealed that duration condition significantly predicted magnitude of error ($t = 16.49$, $p < 0.001$, $r_{partial} = 0.58$), and the consonant presented was also a significant predictor ($t = -6.34$, $p < 0.001$, $r_{partial} = -0.26$). Importantly, language was not a significant predictor when also accounting for duration condition and consonant ($t = -0.17$, $p = 0.87$, $r_{partial} = -0.006$) See Table 2 for detailed statistics.

Table 2. Duration condition and consonant predict error magnitude

Predictor	Partial correlation (*pr*)	*p*-Value
Step 1: $R = 0.05$, *F-change*(1,538) =1.09, *p-change* = 0.30		
Subject	-0.05	0.30
Step 2: $R = 0.61$, *F-change* (3,535) = 104.10, *p-change* < 0.001		
Subject	-0.06	0.06
Language	-0.007	0.92
Duration Condition	0.58	**<0.001**
Consonant	-0.21	**<0.001**

Significant results of added predictors are shown in **bold**.

While both consonant and duration condition were significantly predictive of error magnitude in the above hierarchical regression, it should also be noted that different consonants naturally have different absolute durations. To parse these effects, a second hierarchical regression was conducted using the absolute duration (as opposed to duration condition). This regression included three models. The first model consisted of the individual participant to account for individual differences, the second added the consonant independent variable, and the third model added the absolute duration of the presented token, to explore whether absolute duration accounted for additional variance above and beyond that accounted for by the specific consonant. This analysis (see Table 3) suggests that the absolute duration of tokens is the primary predictor of error magnitude, uniquely accounting for 21% of the variance in error magnitude, with consonant only uniquely accounting for 0.6% of variance.

Table 3. Absolute duration of tokens drives error magnitude

Predictor	Partial correlation (*pr*)	*p*-Value
Step 1: $R = 0.045$, *F-change*(1,538) =1.09, *p-change* = 0.30		
Subject	-0.05	0.30
Step 2: $R = 0.22$, *F-change* (1,537) = 26.78, *p-change* < 0.001		
Subject	-0.05	0.29
Consonant	-0.22	**<0.001**
Step 3: $R = 0.50$, *F-change* (1,536) = 140.73, *p-change* < 0.001		
Subject	-0.03	0.51
Consonant	0.24	**<0.001**
Absolute Duration	0.46	**<0.001**

Significant results of added predictors are shown in **bold**.

This relationship is visually depicted Figure 4A for all trials, and separated by individual consonants in Figure 4B.

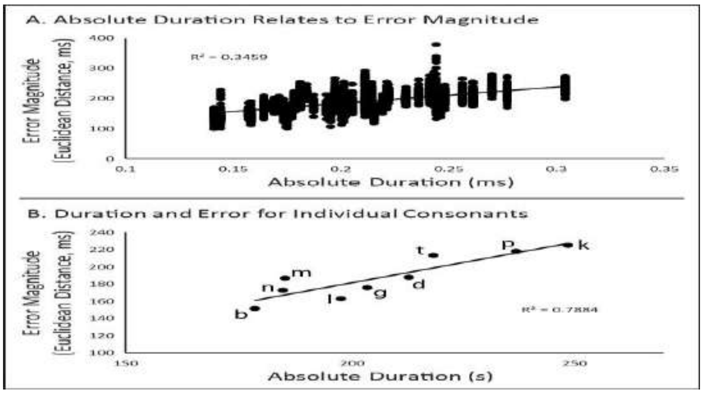


Figure 4. Error rates relative to absolute duration of tokens. Panel A depicts the error magnitude of productions with each individual word token. Panel B depicts error magnitude relative to absolute stimulus duration averaged within each individual consonant. Longer geminate duration in stimuli were strongly predictive of increased error magnitude

Discussion

We had predicted that geminates would be difficult to imitate. This hypothesis was not born out with respect to the rate of geminate production, but it was born out with respect to the degree of target-like productions (accurate length imitation) from an acoustic point of view. The fact that the participants were able to produce geminates, albeit not in a target-like manner, might be due to two factors. First, there is a short-long rhotic contrast in all varieties of Spanish, which may be aiding the participants with the imitation of the geminates sounds in general. The rhotic contrast has traditionally been described as the tap-trill contrast (e.g., Hualde, 2005), but it also exists as a length contrast in some (Andean) varieties of Spanish (Colantoni & Rafat, 2013). It might be that learners are able to apply an existing L1 phonological/phonetic feature to new sounds in an L2 or D2 context. Previous studies have also shown that Spanish-speaking participants are sensitive to length (e.g., Bion, Miyazawa, Kikuchi, & Mazuka, 2013; Escudero, 2001). It is also worth noting that the participants in this study are from mainland Colombia, who reported they had never been in contact with Cuban Spanish before.

However, coastal varieties of Colombian Spanish, such as some varieties from the provinces of Bolivar, Cordoba, and Sucre are very similar to Cuban Spanish, and exhibit gemination. Costal varieties are stigmatized in Colombia and it is possible that our participants were familiar with these varieties or their imitation in media, at least at the perceptual level. We had also predicted that geminates would be equally difficult to imitate in both the L2 and D2 conditions. This hypothesis was born out because the rate of geminate production was the same in both contexts. Moreover, when we looked at the dissimilarity matrices between the L1 and the D2 and L2 productions, although there was a language effect, the results of the regression analysis suggested that length was the main predictor of non-targetlikedness. However, the direction of the effect of length/salience was opposite to what we had initially predicted. This suggests that there might be a trade-off between acoustic salience and ease of imitation in naïve learners. This might be because what may make a sound inherently more salient, may also require more articulatory efforts and/or planning, leading to a difficulty in imitation/production. In the case of acoustic cues, the feature/acoustic attribute 'long' is more difficult to maintain/produce. The hypo- and hyper-articulation patterns of imitation attested here also support this proposal. Whereas there was evidence of both hypo- and hyper-articulation with the short and unaltered conditions (more so with the former than the latter), there was no evidence of hyper-articulation in the long (very salient) condition.

The analysis put forth in this paper sheds light on our understanding of the D2 and L2 speech learning at the onset of acquisition, as we have shown that the participants are able to produce geminates in both contexts, albeit their productions are not acoustically target-like. We have also provided evidence that challenges one of the most important predictions of the Speech Learning Model and highlighted the complexity of the role of salience (length) in L2 and D2 production.

References

Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, *39*(4), 437-456.

Bion, R. A. H, Miyazawa, K., Kikuchi, H., Mazuka, R. (2013). Learning phonemic vowel length from naturalistic recordings of Japanese infant-directed speech. *PLoS ONE*, *8*(2):e51594.

Carlson, K. M., (2011). An acoustic analysis of liquid gemination in the spanish of Havana, Cuba. *Dialectologia: Revista Electrònica*, (6), 1-24.

Celata, C., & Cancila, J. (2010). Phonological attrition and the perception of geminate consonants in the Lucchese community of San Francisco (CA). *International Journal of Bilingualism*, *14*(2), 185-209.

Colantoni, L., & Rafat, Y. (2013). Las consonantes róticas en el español argentino [Rhotic consonants in Argentine Spanish]. In L. Colantoni & C. Rodríguez Louro (eds.), *Perspectivas teóricas y experimentales sobre el español de la Argentina* [Theoretical and experimental perspectives on Argentinian Spanish] (pp. 82-98). Madrid-Frank furt: Iberoamericana/Vervuert.

Colantoni, L., & Steele, J. (2008). Integrating articulatory constraints into models of second language phonological acquisition. *Applied Psycholinguistics*, *29*(3), 489-534.

Escudero, P. (2001). The role of the input in the development of L1 and L2 sound contrasts: Language specific cue weighting for vowels. In A. H.-J. Do, L. Dominguez & A. Johansen (eds.), *Proceedings of the 25th Annual Boston University Conference on Language Development* (BUCLD) (pp. 250-261). Somerville, MA: Cascadilla.

Flege, J. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (ed.), *Speech perception and linguistic perception: Theoretical and methodological issues* (pp.17233-17277). Timonium, MD: York Press.

Hualde, J. (2005). *The sounds of Spanish*, Cambridge, UK: Cambridge University Press.

Khattab, G., & Al-Tamimi, J. (2014). Geminate timing in Lebanese Arabic: the relationship between phonetic timing and phonological structure. *Laboratory Phonology*, *5*(2), 18231-269.

MacLeod, B. (2012). Investigating L2 acquisition of Spanish vocalic sequences. *Studies in Hispanic and Lusophone Linguistics*, *5*(1), 103-148.

Mah, J., & Archibald, J., (2003). Acquisition of L2 length contrasts. In J. M. Liceras, H. Zobl, H. & H. Goodluck (eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*, (pp. 208-212). Somerville, MA: Cascadilla.

Payne, E. M. (2005). Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association*, *35*(2), 153-181.

Rafat, Y. (2011). Orthography-induced transfer in the production of adult novice English speaking learners of Spanish. Unpublished PhD dissertation, University of Toronto.

Rafat, Y. (2015). The interaction of acoustic and orthographic input in the L2 production of assibilated/fricative rhotics. *Applied Psycholinguistics*, *36*(1), 43-64.

Rafat, Y., Mohaghegh, M., & Stevenson, R. (2017). Geminate attrition across three generations of Farsi-English Bilinguals living in Canada: An Acoustic Study. *Ilha do Desterro*, *70*(3), 151-168.

Ruiz-Peña, E., Sevilla, D., & Rafat, Y. (2018). The acquisition of second dialect speech: An acoustic examination of the production of Ecuadorian Spanish assibilated rhotics by Andalusian speakers of Spanish. In E. Babatsouli & D. Ingram (eds.), *Phonology in protolanguage and interlanguage* (pp. 286-315). Sheffield, UK: Equinox Publishing.

Siegel, J. 2010. *Second dialect acquisition*. Cambridge, UK: Cambridge University Press.

Sorianello, P. (2014). Italian geminate consonants in L2 acquisition. Consonant Gemination in First and Second *Language Acquisition* (pp. 25-46). Pisa: Pacini.

To cite this article: Cordero, D., Ruiz-Peña, E., Sierra, E., Stevenson, R., Rafat, Y. (2017). Second dialect and second language imitation of geminates by Colombian Sanish speakers. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 99-105). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Who washed the cup?
Morphological stress deafness. An exploratory study

Syrine Daoussi Díaz, Lorraine Baqué Millet, Marta Estrada Medina
syrine.daoussi@uab.cat, lorraine.baque@uab.cat, marta.estrada@uab.cat

Autonomous University of Barcelona

Abstract. Perceiving stress contrasts is not an easy task for French learners as their mother tongue does not carry a contrastive stress (i.e. stress giving a different semantic value to the lexical unit). Still, in Spanish, such stress contrasts are frequent at a lexical level (*sábana* 'bed sheet' vs. *sabana* 'savannah') but even much more frequent at a morphological level to differentiate first person singular in the present tense from third person singular in the past: *Lavo la taza* 'I wash the cup' vs. *Lavó la taza* 'he/she washed the cup'. Dupoux, Pallier, Sebastián, and Mehler (1997) showed that French learners had difficulties with perceiving these stress contrasts in pseudo-words: stress 'deafness' was also 'persistent' (Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008), i.e. independent of the proficiency level in Spanish of the French participants. This study aimed to assess if stress 'deafness' equally appeared between advanced French learners of Spanish and intermediate learners during an oral comprehension task, involving short texts, sentences and isolated words. The obtained data were analysed by means of linear mixed-effects regression models, in which subjects were introduced as random variables, while L1, type of error (segmental and stress related, lexical or morphological), and item complexity as independent factors. The dependent variable was the Signal Detection Theory measures of sensibility (loglinear A') (Stanislaw & Todorov, 1999). Results show that all groups detected vocalic incoherencies (morphological and lexical) better than stress errors. Within vowel error detection, all groups detected morphological vocalic errors (e.g., *lave la taza* instead of *lavo la taza*) better than lexical vocalic errors (e.g., *levo la taza* instead of *lavo la taza*). Advanced learners perform as well as natives in vocalic detection. If we compare both groups of French learners, there is a significant stress deafness effect, but advanced learners perform significantly better than intermediate students in vowels, but also especially regarding stress errors.

Keywords: prosody, stress deafness, morphology, French, Spanish, oral comprehension

Introduction

Phonological theories underline the strong relationship between perception and production in a native language but also in the case of subsequent language learning. The concept of phonological deafness emerged within the Prague School (Troubetzkoy & Jakobson, 1920-1930). Linguists belonging to this circle started to consider language as a system and began to reflect upon linguistic structures, since each language carries stress specificities, for instance. Spanish and French differ regarding stress pattern, because Spanish stress is free, distinctive and encoded at the lexical level of processing, whereas stress in French is fixed, not-distinctive and generally considered to be computed at the post-lexical level, e.g. *un chat*/*un chat noir/un gros chat noir* 'a cat/a black ca/a big black cat'. There are four existing stress patterns in Spanish (Quilis, 1993): oxytone words (*camión*), paroxytone words (*tragedia*), proparoxytone words (*música*) and superparoxytone ones (*bébetelo*). The acoustic realization of stress in both languages considered differs considerably as stress in French is essentially expressed by a lengthening of the stressed syllable, while in Spanish several parameters are combined, as f0+ duration or f0 +intensity (Llisterri, Machuca, Ríos, & Schwab, 2014). A few years later after its first conception, the metaphor of the phonological deafness materialized into a phonological sieve or 'filter' that would let some elements go, while keeping others on the other side of the barrier (Troubetzkoy, 1967). Phonological deafness has since been transferred to area of stress, leading to the stress 'deafness' concept (Dupoux, Pallier, Sebastián, & Mehler, 1997). In the literature, stress deafness usually refers to difficulties in perceptual processing of specific non-native speech phonemes (Dupoux, 2010). Evidence from these studies points towards the idea that French learners of Spanish show difficulties in perception of lexical stress contrasts in L2 Spanish, leading to 'stress deafness' depending on the cognitive load required by the task and also phonetic variability. This stress deafness takes place in the realm of perception, after acoustic/auditory analysis but before lexical recognition and input to the short term memory buffer. Furthermore, researchers have claimed that it is very robust and resists training, which is the reason why some researchers call it 'persistent stress deafness' (Dupoux, Sebastián-Gallés, Navarrete, 2008). French learners of Spanish are then able to detect 'acoustic' prominence (stress identification, discrimination), but they seem unable to encode stress in phonological representations (lexical processing).

Other studies, however, have showed that French learners are not so 'deaf'. A study on French speakers with no knowledge of Spanish showed that they could identify 87% of the stressed syllables on a written transcription with no stress or punctuation marks, when they heard a spontaneous speech sample (Mora, Courtois, & Cavé, 1997). The authors emphasized that the French learners did not use the same acoustic cues as the Spanish controls. Indeed, they tended to misplace lexical stress to the right, as compared to its original place, as a consequence of the effect of their native language. Other studies have shown that stress deafness is not so 'persistent'. A study on Spanish lexical stress perception by French speakers aimed to see whether the mother tongue may have an influence on Spanish stress perception (Muñoz, Panissal, Billières, & Baqué, 2009), comparing three different groups of French learners of Spanish. The results revealed that learners guessed 83% of the stressed syllables, identified on isolated words and within coherent sentences, and learners with an advanced level of Spanish got better results than intermediate learners or beginners, as they identified all stressed syllables of the bisylabic or trisyllabic words with different stress patterns (oxytone, paroxytone and proparoxytone). Stress deafness was analysed in different tasks, such as stress identification, stress discrimination, lexical decision, and word/pseudoword repetition. Nevertheless, distinctiveness of stress at the morphological (vs. lexical) level has been less explored. Moreover, the effect of stress deafness on oral comprehension tasks is still largely unknown. These considerations regarding stress deafness are also directly related to lexical stress access, as both elements are crucial to better understand the functional architecture of language.

Lexical stress: Retrieved or computed?

In oral comprehension, lexical access is meant to reconstruct a heard message and to extract lexical unities from the continuum of phonemes we are dealing with. Listeners are expected to segment the continuum following their own linguistic rhythm parameters (Cutler, 2012). Traditionally, serial or interactive models are used to explain lexical access, but as researchers have emphasized (e.g., Friederici, 2002), none of them took into account prosodic parameters, and especially stress assignment. If we take a look on how children acquire stress in their mother tongue, they rather seem to learn by rules (or set parameters) than by memorizing stress patterns (Fikkert, 1994). The representation of lexical stress is still under debate and expected to depend on language-specific properties. It seems that for languages in which lexical stress position is fixed (e.g., French), stress pattern need not be stored in the word's phonological representation and can be assigned by default (e.g., Levelt, Roelofs, & Meyer, 1999). Two main theoretical approaches deal with the issue of how stress position is assigned in free-stress languages. While some authors claim that stress pattern of all lexical items is stored in such languages (Butterworth, 1992; Laganaro, Vacheresse, & Frauenfelder, 2002), others argue that regular stress patterns are computed and only the irregular ones are stored and retrieved during word encoding (Colombo, 1992; Levelt *et al*.1999). In addition, an analogical mechanism that assigns stress pattern to a word based on the stress pattern of similar words (e.g., Italian words sharing a particular ending and stress pattern, see Burani, Paizi, & Sulpizio, 2014) has also been considered. An alternative possibility is that stress assignment combines two different processes in free-stress languages: a) the retrieval of a stored representation and b) the computation of the stress pattern based on linguistic (phonological and/or morphological) rules, or statistical distribution (Butterworth, 1992; Laganaro, Vacheresse, & Frauenfelder, 2002).

In Spanish, stress pattern plays a distinctive role. In such languages, lexical stress assignment may occur in several ways: a) access to the lexically stored information (e.g., Spanish *sábana* 'bed sheet' /ˈsabana/ vs. *sabana* 'savannah' /saˈbana/), b) the application of language-specific rules associating phonological or morphological information with stress patterns (e.g., Spanish words ending in a glide are always oxytone, as in *convoy* ‘convoy’ /conˈboj/; every verb-form has a predictable stress: *amo* ‘I love’ /ˈamo/ vs. *amó* ‘he/she loved’ /aˈmo/, *beso* ‘I kiss’ /ˈbeso/ vs. *besó* ‘he/she kissed’ /beˈso/, etc., (Harris, 1983), and/or c) distributional properties (e.g., in Spanish, 90% of nouns, adverbs and adjectives are paroxytones (Harris, 1995) and, more specifically, 95-97% of nouns are either oxytones ending in a consonant or paroxytones ending in a vowel (Alcoba Rueda, 2013; Hualde, 2005).

Objectives and hypothesis

This study aims to assess if stress ‘deafness’ equally appears in advanced French learners of Spanish and intermediate learners during an oral comprehension task of different levels of complexity, as the function condition (target stress pattern), and as compared to vocalic morphological errors and vocalic lexical errors, as the control condition. To assess the persistence of stress ‘deafness’, we compared two groups of French learners (intermediate *vs* advanced).

Method and material

Material

A corpus in Spanish was created (Baqué, L.; Estrada, M.; Daoussi, 2016) of sentences read in acoustically controlled conditions by a male native Spanish phonetician with no language pathology. As Dupoux and collaborators did in earlier studies (Dupoux, Pallier, Sebastián, & Mehler, 1997), we kept the dissociation between phonemes and stress. The current study involves vowel discrimination, unlike in Dupoux and collaborators, who proposed consonant discrimination *versus* stress discrimination. Our corpus consists of three parts with decreasing level of complexity.

The first part, corresponding to 96 short oral texts (see Appendix for examples) were all built on the same model: a first sentence providing the general context in the present or past tense in the 1st or 3rd person; then two short sentences, separated by comas, with the same syllabic structure (verb+direct object); and, finally, another sentence of context. In each sentence, the target verb may have one of these four conditions: i) correct stress pattern, ii) an incorrect stress pattern, iii) incorrect vowel (morphological), iv) incorrect vowel (lexical error). The second part of the test consisted of 48 simple utterances with the same syllabic structure (verb+direct object). Finally, the third part consisted of 48 correct and incorrect verbs in the 1st and 3rd person with incorrect stress pattern or incorrect vowels mixed up together.

Participants

The three groups, matched by gender and education level (tertiary education), were:

- 20 French advanced late learners of Spanish (C1-C2 of CEFRL), 16 females and 4 males, age 40.4 [23-53] in an immersion environment of over three years in Spain (mean=9y.)
- 20 matching native Spanish subjects, 16 females and 4 males, age 38.8 [20-55]
- 38 French intermediate students of Spanish as a foreign language at the University of Toulouse-Le Mirail (France), 31 female and 7 males, age 21.16 [18-55]

All of them reported suffering from no hearing or language disorders. They signed the written consent before taking part in the study.

Procedure

The advanced learners took the perception test online on the Labguistic platform (Ménétrey & Schwab, 2014), where the test was hosted. The intermediate learners took the test in a computer room at the University of Toulouse Le Mirail (France). All participants had a short training session to ensure that they understood what was expected of them, i.e. if they were incoherencies of any type (lexical or morphological) in the sentences they heard. The perception test was online and took around 40 minutes.

Data analysis

For each comparison, we carried out a means of sensibility (loglinear A') (Stanislaw & Todorov, 1999) with linear mixed-effects regression models, in which subjects were introduced as random variables. L1, type of error (segmental and stress related, lexical or morphological) and item complexity appeared as independent factors.

Results

There is a main effect of Group (Wald-$\chi2$ (2) = 167.79, p = .000) and Type of error (stress_morphological vs. vocalic_morphological, and vocalic_lexical) (Wald-(2) = 257.09, p = .000). There is an interaction effect between Group and Type of error (Wald-$\chi2$ (4) = 88.70, p = .000).

Post-hoc tests show that native Spanish controls (henceforward ES) performed better in all conditions. All three groups showed less sensitivity to stress errors than to morphological and lexical vocalic errors. Intermediate learners (henceforward B1-B2) presented poorer sensitivity to lexical vowel, as compared to morphological vowel and to morphological stress vs. which is always significantly lower than advanced learners (henceforward ADV).

-VOY_GRAM: loglinear A' mean is respectively 0.94893 (ES), 0.94079 (ADV), 0.81806 (B1-B2) p=.000

-VOY_LEX: loglinear A' mean is respectively 0.94186 (ES), 0.93842 (ADV), 0.76725 (B1-B2) p=.000

-STRESS_MORPHO: loglinear A' mean is respectively 0.89667 (ES), 0.73277 (ADV), 0.56956 (B1-B2) p=.000

Figure 1. Effect of error type on morphological error detection: Group*Error Type

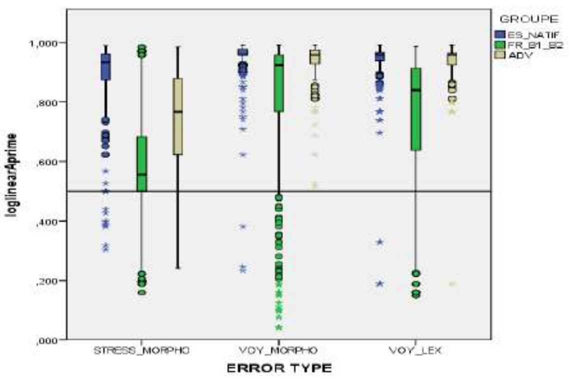
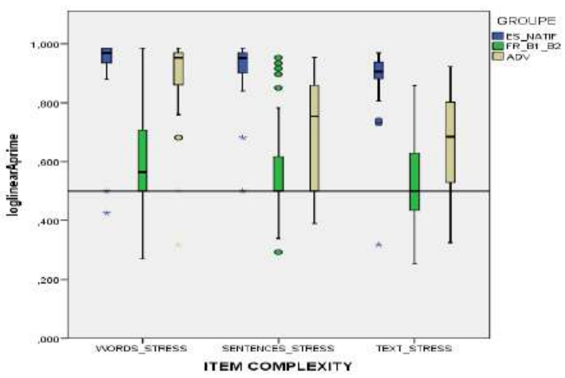


Figure 2. Effect of complexity on morphological error detection: Group *Complexity (Texts, Sentences, Words)



These results show that, regarding error detection on morphological vowels, advanced learners perform almost like the native controls, 0.94079 (ADV) and 0.94893 (ES), p=.000, respectively. Regarding error detection on lexical vowels, advanced learners also perform the native controls: 0.93842 (ADV), and 0.94186 (ES), p=.000, respectively. Regarding error detection on morphological stress, advanced learners are closer to native controls than intermediate learners, who showed almost no sensitivity to stress contrasts: respectively 0.73277 (ADV), 0.89667 (ES), 0.56956 (B1-B2) p=.000.

There is a main effect of Group (Wald-$\chi2$ (2) = 242.750, p = .000) and Item Complexity Wald-$\chi2$ (2) = 39.291, p = .000). There is an interaction effect between Group and Item Complexity (Wald-$\chi2$ (4) = 15.418, p = .004).

Post-hoc tests show that native Spanish controls (henceforward ES) performed better in all conditions.

-WORDS_STRESS: loglinear A' mean is respectively 0.93769 (ES), 0.86921 (ADV), 0.61711 (B1-B2) p=.000

-SENTENCES_STRESS: loglinear A' mean is respectively 0.91980 (ES), 0.71668 (ADV), 0.55499 (B1-B2) p=.000

-TEXTS STRESS: loglinear A' mean is respectively 0.88559 (ES), 0.68211 (ADV), 0.53742 (B1-B2) p=.000

Figure 3. Effect of error position on morphological error detection: Group*Position (1st Vs 2nd Verb)

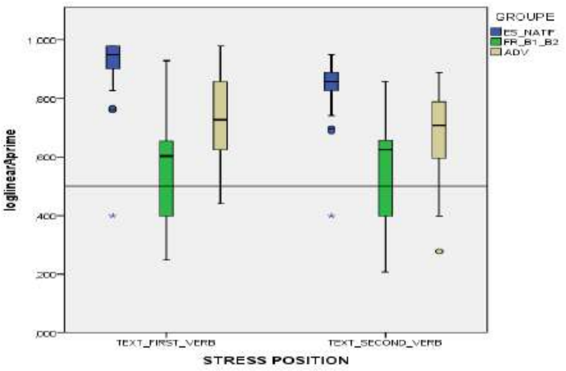
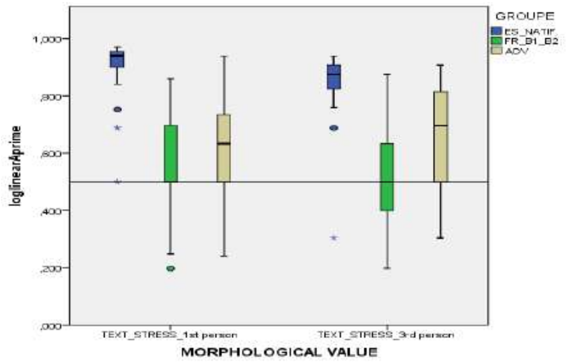


Figure 4. Effect of morphological value on morphological error detection: Group *Morphological Value (Present 1st Person vs Past 3rd Person)



These results show that for all three groups of learners, item complexity is significantly relevant, i.e. they show more sensitivity to less complex item (words). Their sensitivity decreases on sentence errors, and decreases even more in texts. However, results also shed light on the fact that sensitivity to item complexity is gradual for ES, while there is a threshold with higher detection between words *vs* sentences and texts for both groups of French learners.

There is a main effect of Group (Wald-$\chi 2$ (2) = 204.396, p = .000) and Stress Position (Text_mot1_text_mot2) Wald-$\chi 2$ (1) = 12.679, p = .000). There is an interaction effect between Group and Stress Position (Wald-$\chi 2$ (2) = 11.096, p = .004).

Post-hoc tests show that native Spanish controls performed better in all conditions, although they detected errors on the first verb better than on the second. Advanced learners (henceforward ADV) showed less sensitivity than ES but parallel evolution, since errors are perceived on the first verb better than on the second verb. Intermediate learners (henceforward B1-B2) showed no sensitivity to any of the possible error position.

-TEXT_FIRST_VERB: loglinear A' mean is respectively 0.91803 (ES), p=.000, 0.72086 (ADV), p=.000, 0.55522 (B1-B2) p=.000

-TEXT_SECOND VERB: loglinear A' mean is respectively 0.84782 (ES), p=.000, 0.67551 (ADV) p=.000, 0.55841 (B1-B2) p=.000

The results show that there is a difference in the two groups of French learners: B1-B2 showed no sensitivity at all to error position, while ADV detected errors on the first position more easily, just like ES did.

There is a main effect of Group (Wald-$\chi 2$ (2) = 252.447, p = .000) and Morphological Value Wald-$\chi 2$ (1) = 0.799, p = .000), and there is an interaction effect between Group and Morphological Value (Wald-$\chi 2$ (2) = 6.071, p = .048).

Post-hoc tests show that native Spanish controls performed better in all conditions, although they detected errors on paroxytone stress pattern (text_stress_1st person) better. Advanced l earners (henceforward ADV) showed less sensitivity than ES and showed more sensitivity to errors with an oxytone pattern (text stress_3rd person). Intermediate learners showed no sensitivity to any of both patterns presented.

-TEXT_STRESS_1st person: loglinear A' mean is respectively 0.90022 (ES), p=.000; 0.62635 (ADV) p=.000, 0.55135(B1-B2) p=.000

-TEXT_STRESS 3rd person: loglinear A' mean is respectively 0.84482 (ES), p=.000; 0.67710 (ADV) p=.000, 0.51010 (B1-B2), p=.000

The results show that intermediate learners showed no sensitivity to oxytone or paroxytone stress patterns. ES showed more sensitivity when the paroxytone stress pattern is violated, while ADV showed more sensitivity when the oxytone stress pattern is violated.

Conclusion

This preliminary paper on stress deafness, regarding morphological items, allows researchers to agree with Dupoux and colleagues who previously studied non-words (1997), that perceiving stress errors is more difficult that perceiving vocalic errors. Both groups of French learners in this study showed stress 'deafness' in morphological stress contrasts. Nevertheless, this study emphasizes the fact that the Spanish proficiency level of French learners seems to relate to the vocalic and stress perception of morphological contrasts. But to gain confidence, results should be replicated in one series for both lexical and morphological stress errors.

Acknowledgements

This study was funded by the Spanish Ministry of Economy and Competitiveness (COGNIPROS Project (FF12013-40419-P) with Lorraine Baqué and Marta Estrada as Principal Investigators.

References

Baqué, L., Estrada, M., Daoussi, S. (2016). *Cognipros Corpus (morphological stress)*.

Butterworth, B. (1992). Disorders of phonological encoding. *Cognition*, *42*, 261-286.

Colombo, L. (1992). Lexical stress effect and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 987-1003.

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. *Native Listening: Language Experience and the Recognition of Spoken Words*. Cambridge: MA: MIT Press.

Dupoux, E. (2010). *Phonological 'deafness'. Summary of research.* Retrieved from http://www.lscp.net/persons /dupoux/presentation/Dupoux_summary_phonological_deafnesses.pdf

Dupoux, E., Pallier, C., Sebastián, N., Mehler, J. (1997). A destressing 'Deafness' in French? *Journal of Memory and Language*, *36*, 406-421. Retrieved from https://pdfs.semanticscholar.org/ 0c8f/f9ce3660dae2d4aa58ebb 89b 390a5ea4f0c0.pdf

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition*, *106*(2), 682-706.

Fikkert, P. (1994). *On the acquisition of prosodic structure*. *Radboud Repository*. Dordrecht: lCG Printing.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78-84.

Laganaro, M., Vacheresse, F., & Frauenfelder, U. H. (2002). Selective impairment of lexical stress assignment in an Italian-speaking aphasic patient. *Brain and Language*, *81*(1-3), 601-609.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences*, *22*(1), 1-38.

Llisterri, J., Machuca, M. J., Ríos, A., & Schwab, S. (2014). El acento léxico en contexto: Datos acústicos. *Fonética Experimental, Educación Superior E Investigación. Adquisición Y Aprendizaje de lenguas/Español Como Lengua Extranjera.*, *1*, 357-376.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*(1), 137-149. Retrieved from http://people.brandeis.edu/~sekuler /stanislawTodorov1999.pdf

Troubetzkoy, N. S. (1967). *Principes de phonologie* Paris: Klincksieck.

Appendix

Corpus sentences. Examples.

Short texts

1. → coherent

Yo, cada mañana, desayuno en casa. Lavo la taza, miro la tele, y me voy a trabajar.
‘Every morning, I have breakfast at home. I wash the cup, I watch TV, and I go to work’

2. → incorrect stress pattern creating morphological incoherence

Yo, cada mañana, desayuno en casa. Lavó la taza, miro la tele, y me voy a trabajar
‘Every morning, I have breakfast at home. He washed the cup, I watch TV, and I go to work’

3. → incorrect vowel creating lexical (semantic) incoherence

Yo, cada mañana, desayuno en casa. Levo la taza, miro la tele, y me voy a trabajar. ‘Every morning, I have breakfast at home. I break out [anchor] the cup, I watch TV, and I go to work’

4. → incorrect vowel creating a morphological incoherence

Yo, cada mañana, desayuno en casa. Lave la taza, miro la tele, y me voy a trabajar.
‘Every morning, I have breakfast at home. Wash [imperative] the cup, I watch TV, and I go to work’

Sentences

1. Cené la sopa ‘I ate soup for dinner’ → coherent (3rd person, past)
2. Levo la taza ‘I break out [anchor] the cup’ → incorrect vowel: lexical incoherence
3. Ffumo la pipa ‘I smoke the pipe’ → coherent (1st person, present)
4. Cene la sopa ‘eat[imperative] soup for dinner’ → incorrect vowel: morphological incoherence

Isolated words

1. saco ‘I take out’ → coherent (1st person, present)
2. dejó ‘he/she left’ → coherent (3rd person, past)
3. saque ‘take out’ → coherent (imperative, 3d person)
4. sicó ‘he took out’ → incorrect vowel, lexical (semantic) incoherence

To cite this article: Daoussi, S., Baqué, L., & Estrada, M. (2017). Who washed the cup? Morphological stress deafness in French learners. An exploratory study. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 106-112). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

A real time study of contact-induced language change in Frisian relative pronouns

Jelske Dijkstra[1], Wilbert Heeringa[1], Emre Yılmaz[2], Henk van den Heuvel[2], David van Leeuwen[2], Hans Van de Velde[1]

jdijkstra@fryske-akademy.nl, w.heeringa@fryske-akademy.nl, e.yilmaz@let.ru.nl, h.vandenheuvel@let.ru.nl, d.vanleeuwen@let.ru.nl, hvandevelde@fryske-akademy.nl

[1]Fryske Akademy, Leeuwarden, [2]Radboud University, Nijmegen

Abstract. Many minority languages are subject to linguistic interferences from a more prestigious language as, for example, the country's majority language. This is also the case with (West-)Frisian, spoken in the province of Fryslân in the north of the Netherlands. The current real-time study investigates the influence of Dutch on the use of Frisian relative pronouns in the semi- and non-scripted speech of speakers of Frisian in a corpus of 18.5 hours of audio fragments from the regional broadcaster *Omrop Fryslân*. The audio fragments are retrieved from various radio programs broadcast between 1966 and 2015. Each fragment lasts approximately five minutes. The current study focuses on t-deletion in the Frisian relative pronouns *dy't* 'who/that', *dêr't* 'where' and *wêr't* 'where'. The data are analyzed using cumulative link mixed models to check the influence of the factors: year of broadcast, speaking rate and age of the speakers on t-deletion. Results show that t-deletion occurs more often in recent broadcasts than in older ones, but speaking rate and age of the speakers are not significant factors. These outcomes demonstrate that this language change in Frisian has recently been set in motion and that it seems to follow a generational change pattern.

Keywords: language change, minority language, majority language, relative pronouns, t-deletion, generational change

Introduction

Minority languages and regional varieties often face pressure from a more prestigious language, usually a majority language spoken in that region (e.g., Padovan, Tomaselli, Bergstra, Corver, Etxepare, & Dol, 2016; Sankoff, 2001). Such situations of language contact usually result in the borrowing of lexical elements and interference of grammatical elements from the more prestigious language into the other variety. When enough speakers adopt these borrowings and interferences in their speech, these 'foreign' elements gradually become part of the minority language, resulting in language change. Language change can take on different patterns, depending on the stability of a language feature within the individual and within the community as a whole (Evans Wagner, 2012; Labov, 1994; Sankoff, 2005). Four types of change can be distinguished: (1) in generational change, a language feature stays stable during the adult life of individuals but changes over time within the community as a result of the fact that new generations show an increase in the adoption of that feature; (2) in age-grading, a language feature is instable in the speech of the individuals but stable over time within the community as a whole; (3) in communal change, both individuals and the entire community take over the innovation rapidly and simultaneously; (4) in lifespan change (Sankoff, 2005), a certain feature changes within individuals in the same direction as the change that takes place within the community over time. It is, however, rather difficult to make an explicit distinction between the different patterns of language change. For example, real-time data and information about social factors are required to differentiate age-grading from generational change or life span change. And in many cases, this information is lacking, especially when archival material is used (Evans Wagner, 2012).

As mentioned, the direction of borrowings and interferences is often from a standard or majority language to a minority language or regional variety. This is for example the case of Cimbrian, a Germanic minority language spoken in the province of Trento and the Veneto region in the northeast of Italy. One of the two complementizers in relative clauses used in Cimbrian is borrowed from Italian and is gradually spreading into the Cimbrian language (Padovan *et al*., 2016). Other examples are the Maaseiks dialect in Belgium, Susters in the Netherlands, and Waldfeuchts in Germany. These three closely related dialects have undergone change on the lexical, phonological and morphological levels, due to the influence of three different overarching standard varieties i.e. Belgian Dutch, Netherlandic Dutch and Standard German respectively (Gerritsen, 1999).

West-Frisian, a minority language spoken in the province of Fryslân in the Netherlands, is also changing under the influence of Dutch. Other varieties of Frisian are spoken in Germany, i.e. North Frisian and Saater Frisian. Whenever Frisian is mentioned in this paper, we refer to the West-Frisian variety. Like Dutch, Frisian is a West-Germanic language. All speakers of Frisian are bilingual Frisian/Dutch, since Dutch is the main language used in education and most language domains (Gorter & Jonkman, 1995). About 55% of the inhabitants of Fryslân are native speakers of Frisian and another 20% speak Frisian as a second language (Gorter & Jonkman, 1995; Provinsje Fryslân, 2015).

The current paper investigates a change in the Frisian relative pronouns *dy't* 'who/that', *dêr't* 'where' and *wêr't* 'where' under the influence of Dutch. The study focusses on deletion of word-final /t/ in these three relative pronouns. The question is whether there is an increase in t-deletion in *dy't*, *dêr't* and *wêr't* in Frisian speech of recent times. So far, no quantitative study has been conducted on this subject.

Frisian and Dutch relative pronouns *who* and *where*

Table 1 presents the Frisian relative pronouns *dy't* 'who/that', *dêr't* 'where' and *wêr't* 'where' and their Dutch counterparts.

Table 1. Relative pronouns in Frisian and Dutch

Frisian			Dutch			English translation
dy't	[dit]	*de man/frou dy't in boek lêst*	*die*	[di]	*de man/vrouw die een boek leest*	'the man/woman who is reading a book'
		de auto's dy't foarby ride			*de auto's die voorbij rijden*	'the cars that drive past'
dêr't	[dɛːt]	*de stêd dêr't er no wennet*	*waar*	[ʋaːr]	*de stad waar hij nu woont*	'the city he currently lives in'
wêr't	[ʋɛːt]	*wêr't de kaai is, wit ik net*	*waar*	[ʋaːr]	*waar de sleutel is, weet ik niet*	'I do not know where the key is'

The Frisian relative pronoun *dy't* and its Dutch counterpart *die* are used with masculine, feminine, and plural antecedents. When the antecedent is considered a location, the relative adverb *dêr't* or *wêr't* is used. If there is an antecedent, *dêr't* is used in Frisian. In free relative clauses, *wêr't* is obligatory (see Table 1). They have the same Dutch counterpart, namely *waar*.

As Table 1 shows, *dy't*, *dêr't* and *wêr't* all exhibit word-final /t/, whereas their Dutch counterparts do not. This clitic does not only appear in these relative pronouns but also in several other conjunctions heading a subordinate clause (Taalportaal, 2017). In the case of *dy't*, *dêr't* and *wêr't*, the complementizer *'t* is a marker that the pronoun *dy*, or the adverb *dêr* or *wêr* is relativized. Several sources claim it represents a reduced form of the conjunction *dat* [dɑt] 'that' (Hoekstra, 2002; Van Coetsem, 1960; Van der Meer, 1991). However, Van der Woude (1960) questions this, since he found only a few conjunctions combined with *dat* in Middle Frisian sources. The form *'t* is a relatively new phenomenon in the Frisian language. It is mentioned for the first time in *dy't* as an option next to *dy* in a descriptive Frisian grammar at the end of the 19th century (Van Blom, 1889). Although Brouwer (1959) noted an increase in the use of the clitic /t/ in the conjunctions that introduce subordinate sentences, a couple of decennia later Van der Meer (1991) observed the contrary: 'Due to Dutch influence the phenomenon [presence of clitic /t/] seems to be receding again, not only among younger generations' (p.44). Both sources are based only on observations or written sources, rather than on empirical studies based on live speech corpora.

Method

This section describes the methodology used to answer the research question. It starts with a description of the corpus that is used in this paper. Next, the procedure and analysis are described.

FAME! Speech Corpus

The data were selected from the FAME! Speech Corpus, which is a corpus used for training automatic speech recognition (see also www.fame.frl). This corpus comprises 203 audio fragments with a duration of approximately five minutes each (total 18.5 hours). The audio fragments are randomly selected from various radio programs in the archive of the regional broadcaster *Omrop Fryslân*, e.g., discussion programs, quizzes, news items, etc., and a variety of topics, such as agriculture, nature, life stories, sports, culture, literature, etc. The fragments were broadcast between 1966 and 2015.

All audio fragments of the FAME! Speech Corpus were manually transcribed by two native speakers of Frisian and checked afterwards by the first author, a native speaker of Frisian too. Speech of non-native speakers was discarded, just as speech that was annotated as scripted radio, i.e. radio where one reads a previously written message aloud. Only non- and semi-scripted speech was used. In semi-scripted speech, the speaker uses prewritten catchwords or phrases. Repetitions and revisions were discarded as well. Further, speech contexts where it was unclear whether /t/ was pronounced or not, such as when the target word was followed by a t-initial word, e.g., *dy't tydlik* [ditidlək] 'that temporarely', or in case of second person singular when the target word has an inflectional ending in *–st* e.g,. *dy'tst hast* [disthɑst] 'that-you have', were left out of the data. Table 2 shows the characteristics of the final data set for /t/.

Table 2. Characteristics of data set FAME! Speech Corpus

# tokens (# speakers)	*dy't*	351	(133)
	dêr't	27	(18)
	wêr't	111	(73)
# speakers (# unique speakers)		158	(120)

The relative pronoun *dy't* shows the most tokens, i.e. 351 tokens used by 133 speakers. It must be mentioned, however, that some speakers appear multiple times in the corpus. The high number of tokens for the relative pronoun *dy't* is in contrast with the occurrence of the relative adverb *dêr't*, which only has 27 tokens in the data, spoken by 18 speakers. For *wêr't*, 111 tokens were found, used by 73 speakers. A total of 158 speakers (120 unique speakers) used one of these three relative pronouns at least once.

Procedure

In each audio fragment the number of times the target word occurred was counted. This means that both the number of deletions and realizations of /t/ were counted. Further, the following segment was coded, to be able to control for traditional linguistic factors influencing t-deletion. Speaking rate (Laver, 1994) was automatically measured for each speaker within each fragment. Additionally, age was calculated by retrieving the year of birth of all speakers. For 70% of the speakers, the year of birth could be traced by means of searches on the internet. The year of birth was estimated for the remaining 30% of the speakers, mainly on the basis of voice characteristics and contextual indicators. The mean age of the speakers was 49.9 years (SD 14.3 years).

Analyses

The data were analyzed using the Cumulative Link Mixed Model (CLMM) in R (The R Foundation for Statistical Computing, http://CRAN.R-project.org) by applying the clmm function in the Ordinal package (Christensen, 2015). The factor t-deletion with 0 indicating t-deletion and 1 indicating t-realization was entered as ordered dependent variable. The factor speaker was included as random intercept. Some speakers appeared multiple times in the corpus or used at least one of the relative pronouns multiple times. The linguistic context following the pronoun was also included as a random intercept, so the data was controlled for this factor. Pronoun was added as random slope of linguistic context which improved the quality of the model (measured by the Akaike Information Criterion - AIC) further. This random slope controls for the fact that the influence of linguistic context may be different per pronoun. The initial model contained all fixed factors under investigation, i.e. year of broadcast, age of speakers and speech rate, and its possible interactions. By considering all possible combinations of the fixed factors we found the optimal model with year of broadcast, age and the interaction between broadcasting year and age as fixed factors. Analyzing this model we obtained the results that are shown in Table 3.

Table 3. Optimal model for t-deletion in *dy't, dêr't* and *wêr't*

	Estimate	Standard error	Z value	Pr(>|z|)
year of broadcast	-0.68	0.32	-2.14	p<.05
age	0.21	0.33	0.65	n.s.
year of broadcast * age	0.96	0.34	2.85	p<.01

Results

In Table 3, we find that the year of broadcast is a significant factor in t-deletion in the three relative pronouns. The negative estimate score means that the more recent the audio fragment dates are, the more word-final /t/ is deleted in *dy't, dêr't* and *wêr't*. This is confirmed by Figure 1. The figure shows a decrease of /t/'s being pronounced over the three time periods. Note that neither a 100 percent realization nor a 100 percent deletion of /t/ is reached in either time period. And even in recent broadcasts, /t/ is still realized over half of the times in these three relative pronouns.

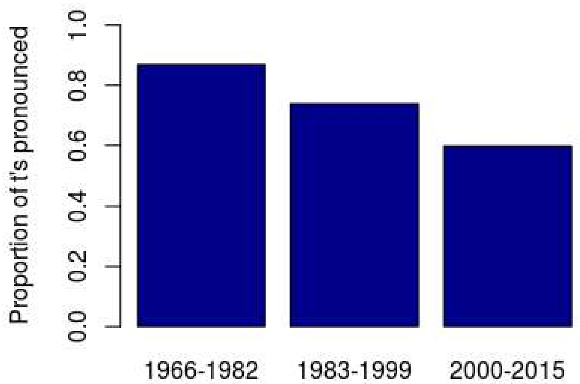


Figure 1. Histogram showing percentages of t-realizations dy't, dêr't and wêr't over time

Speaking rate does not play a significant role in the process of t-deletion in these contexts. The age of the speaker is present in the optimal model, but not as a significant factor. Age is only significant in interaction with broadcasting year. Figure 2 presents the interaction between these two factors. For reasons of clarity, both age and year of broadcast are divided into ranges in this figure. The figure shows the trend that the proportion of /t/'s that are pronounced decreases with age in the period 1966-1982. For the period 2000-2015, we find the opposite trend. For the period 1983-1999, we see that the proportion of t-deletions shows a slight increase with age. We also find that the differences between the three time periods are large for the youngest age group but small for the oldest age group. Table 4 in Appendix 1 presents the percentages of t-realizations for each relative pronoun per broadcasting period and age group.

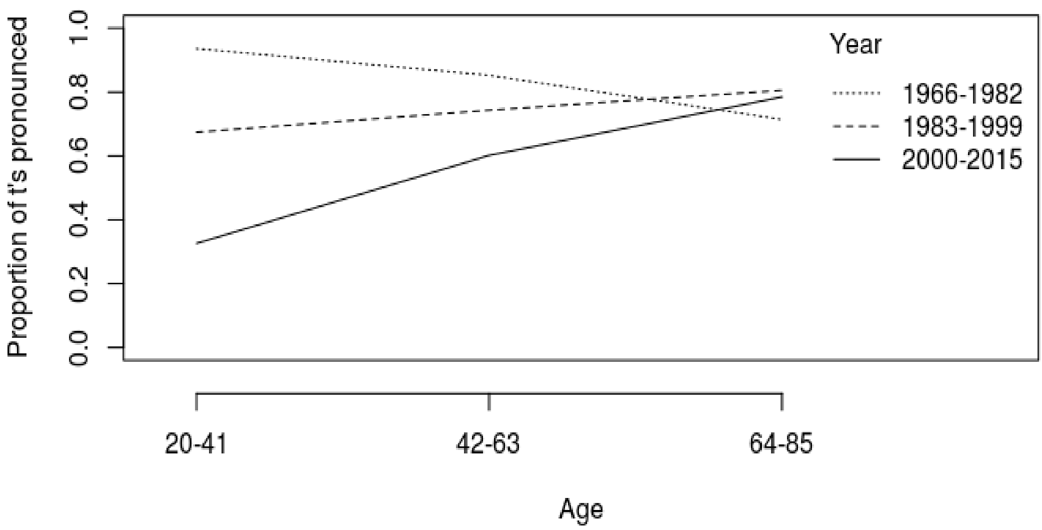


Figure 2. Interaction between year of broadcast and age

Discussion

The year of broadcast is a significant factor in the process of t-deletion in the relative pronouns *dy't*, *dêr't* and *wêr't*. The increase in t-deletion confirms the observation of Van der Meer (1991) that word-final /t/ was receding in several conjunctions. The relatively low number of tokens for *dêr't* in the corpus can be explained by the fact that *dêr't* is often substituted by *wêr't* under the influence of Dutch (De Haan, 2001; Taalportaal, 2017).

Radio speech from the 1950s-1970s is more formal than speech in later years (see also Van de Velde, 1996). One of the characteristics of the deformalization of radio speech in recent years is a faster speaking rate. However, this paper shows that despite the informal character of recent radio speech, speaking rate did not elicit significantly more deletions of word-final /t/ in these three relative pronouns.

The age of the speakers was not a significant factor, only in interaction with broadcasting year. T-deletion occurred among speakers from all age groups present in the corpus. In the older broadcasts, it seems that the older the speakers were, the more t-deletion they used. In more recent broadcasts, we see the opposite, namely that the younger speakers used more t-deletion compared to the older speakers. This pattern can be explained as follows. At the end of the 19th century, *'t* was introduced as a new form in Frisian. In the 1960s and 1970s, it was still a change in progress that was led by younger speakers. Then, around the 1980s, the change is reversed: younger speakers start deleting /t/ in *dy't*, *dêr't* and *wêr't* and take the lead in this process in the 21st century. This phenomenon has been studied in more detail in Dijkstra, Heeringa, Yılmaz, Heuvel, & Leeuwen (2017).

Because the data concerned archive material from a radio station, sociolinguistic information of speakers was absent in many cases. The FAME! Speech Corpus with 18.5 hours of speech was initially not collected for the purpose of language variation research. It only comprises of 489 tokens of *dy't*, *dêr't* or *wêr't*, which is a rather low number for investigating which particular linguistic contexts elicits t-deletion or not. The relatively small corpus entailed that it was only possible to control for these contexts, not to look more deeply into this factor.

Conclusion

The current paper presents language change in progress, i.e. t-deletion in the Frisian relative pronouns *dy't*, *dêr't* and *wêr't*. The study was conducted on an 18.5-hours long corpus of 203 audio fragments randomly selected from different radio programs broadcasted by the regional broadcaster *Omrop Fryslân.*

Results show there is a significant increase of t-deletion in *dy't, dêr't* and *wêr't* over time. The year of broadcast is a significant factor, whereas the age of the speakers is not. In other words, the community seems to gradually adopt t-deletion in these three relative pronouns, and this phenomenon occurs among speakers of all ages. This indicates that a generational change (e.g., Evans Wagner, 2012; Labov, 1994; Sankoff, 2005) is in progress here. It should, however, be investigated further in a large-scale panel study whether this change indeed involves a generational change.

The database that was used in the current paper, i.e. the FAME! Speech Corpus was developed to train an automatic speech recognizer for Frisian. With this speech recognizer, the entire radio corpus of *Omrop Fryslân* will be disclosed in the nearby future resulting in a speech database of more than 2600 hours of radio speech. Preferably, the study is repeated with that large corpus to see which linguistic contexts following the pronoun influence this process.

Acknowledgments

The FAME! Project is funded by the Netherlands Organisation for Scientific Research NWO, Project 314-99-119 (Frisian Audio Mining Enterprise), which is gratefully acknowledged. We would further like to thank Maaike Andringa and Sigrid Kingma for meticulously annotating the radio fragments in the FAME! Speech Corpus.

References

Brouwer, J. H. (1959). Mei té of sûnder té. *De Pompebledden*, *30*, 62-64.

Christensen, R. H. B. (2015). Ordinal – Regression Models for Ordinal Data. R package version 2015.6-28. Retrieved from http://www.cran.r-project.org/package=ordinal/

De Haan, G. J. (2001). *Grammar of modern West Frisian*. Groningen, Netherlands: University of Groningen.

Dijkstra, J., Heeringa, W., Yılmaz, E., Heuvel, H. van den, & Leeuwen, D. van. (2017). Language change caught in the act: a case-study of t-deletion in Frisian relative pronouns. Manuscript submitted for publication.

Evans Wagner, S. (2012). Age grading in sociolinguistic theory. *Language and Linguistics Compass*, *6*(6), 371-382.

Gerritsen, M. (1999). Divergence of dialects in a linguistic laboratory near the Belgian – Dutch – German border: Similar dialects under the influence of different standard languages. *Language Variation and Change*, *11*, 43-65.

Gorter, D., & Jonkman, R. J. (1995). *Taal yn Fryslân. Op 'e nij besjoen*. Leeuwarden, Netherlands: Fryske Akademy.

Hoekstra, J. (2002). Relativisation in Frisian. In P. Poussa (ed.), *Relativisation on the North Sea littoral* (pp. 63-76). Muenchen: Lincom Europa.

Labov, W. (1994). *Principles of linguistic change: Internal factors*. Oxford, UK: Blackwell.

Laver, J. (1994). *Principles of phonetics*. Cambridge, UK: Cambridge University Press.

Padovan, A., Tomaselli, A., Bergstra, M., Corver, N., Etxepare, R., & Dold, S. (2016). Minority languages in language contact situations: three case studies on language change. *Us Wurk*, *65*, 146-174.

Provinsje Fryslân (2015). *De Fryske taalatlas. Fryske taal yn byld*. Leeuwarden, Netherlands: Provinsje Fryslân.

Sankoff, G. (2001). Linguistic outcomes of language contact. In P. Trudgill, J. Chambers & N. Schilling-Estes (eds.), *Handbook of Sociolinguistics* (pp. 638-668). Oxford, UK: Blackwell.

Sankoff, G. (2005). Cross-sectional and longitudinal studies in sociolinguistics. In P. Trudgill (ed.), *Sociolinguistics: An international handbook of the science of language and society* (pp. 1003-1013). Berlin: Mouton de Gruyter.

Taalportaal (2017). Relative pronouns. Retrieved from http://www.taalportaal.org/taalportaal/topic/pid/topic-13998813311730000
Van Blom, P. (1889). *Beknopte Friesche spraakkunst voor den tegenwoordigen tijd*. Joure: Zijlstra.
Van Coetsem, F. (1960). De Friese relativa en conjuncties met 't. *Fryske Stúdzjes*, 327-334.
Van de Velde, H. (1996). *Variatie en verandering in het gesproken Standaard-Nederlands (1935-1993).* (Doctoral dissertation) University of Nijmegen.
Van der Meer, G. (1991). The subclause signal ‘t in Frisian. Its origin and function. *Leuvense Bijdragen*, *80*, 43-59.
Van der Woude, G. (1960). Oer it gebrûk fan 't by bynwurden. *Fryske Stúdzjes*, 335-343.

Appendix 1

Table 4. Percentages of t-realizations for *dy't*, *dêr't* and *wêr't* per broadcasting period and age group

Broadcasting period	Age group	*dy't*	*dêr't*	*wêr't*
1966-1982	20-41	97.30	66.67	85.71
	42-63	86.96	87.50	57.14
	64-85	70.00	0.00	75.00
1983-1999	20-41	80.65	100.00	27.27
	42-63	78.85	100.00	42.86
	64-85	72.00	100.00	100.00
2000-2015	20-41	27.59	100.00	33.33
	42-63	62.50	100.00	53.33
	64-85	79.59	50.00	78.57

To cite this article: Dijkstra, J. Heeringa, W., Yilmaz, E., Heuvel, H. van den, Leeuwen, D. van, & Velde, H. van de (2017). A real time study of contact-induced change in Frisian relative pronouns. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 113-119). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

L1 Farsi attrition in contact with L2 Canadian English: A focus on rhotics and stress pattern

Nasim Fakoornia
nfakoorn@uwo.ca

University of Western Ontario

Abstract. This study addresses first language (L1) phonetic attrition, and second language (L2) phonetic acquisition of a 9-year-old Farsi-English bilingual child. The study aimed to investigate how the manner of articulation (MOA) of the phoneme /r/, and the two correlates of stress, F0 peak and syllable duration, may change in both the L1 and L2 speech of a Farsi-speaking newcomer to Anglophone Canada within a period of one month. The experimental procedure included a picture-naming task and was carried out in two sessions, that included Farsi and English. It must be taken into account that in English, approximants are the most dominant allophones of /r/, whereas rhotics in Farsi have a number of different allophonic realizations in different positions. Regarding the stress pattern in English, stress in bi-syllabic nouns is often on the first syllable, whereas in Farsi stress in nouns is on the final syllable. In this regard, results of the acoustic analysis revealed that in the second session the number of approximants in Farsi in most positions was increased, providing evidence for L1 attrition in Farsi. On the other hand, in the second session the stress in Farsi words was misplaced on the first syllable. Yet, as opposed to syllable duration, the F0 peak was not a consistent factor in determining the change in stress pattern. The measurement of rhotics in English showed that the majority of rhotics were approximants in both sessions. Moreover, accuracy in producing the duration of syllables and the location of F0 peak in English increased, resulting in producing more English-like tokens. Thus, it was concluded that the child was acquiring English L2 phonology. The findings of this study are novel and contribute to our understanding of attrition and L2 acquisition in child phonology.

Keywords: L1 Farsi attrition, L2 English acquisition, manner of articulation of rhotics, correlates of stress

Introduction

Although a rich body of literature has been devoted to the phenomenon of attrition, little has been done on attrition in child or heritage phonology (e.g., Rafat, Mohaghegh, & Stevenson, 2017).

This is an acoustic study of first language (L1) phonetic attrition and second language (L2) phonetic acquisition of a 9-year-old Farsi-English bilingual child. The aim of this research was to investigate how the manner of articulation (MOA) of phoneme /r/ may change in both the L1 and L2 speech of a Farsi-speaking newcomer child to Anglophone Canada within a period of one month. Furthermore, considering the difference of stress patterns (SP) in Farsi and English, the other goal of the study was to examine whether two correlates of stress, F0 peak and syllable duration in L1 and L2, bi-directionally affect each other at the word level or not. In this respect, this research attempts to answer the following questions:

1. Does attrition occur with respect to manner of articulation and stress pattern in the child's Farsi?
2. Does the child acquire manner of articulation and stress pattern in English?

Language attrition is the loss of or change of different linguistic levels of the L1 of speakers who have changed their linguistic environment and language habits (Schmid, 2011). According to Major (1992), the more one is proficient in L2, the more loss occurs in his L1. Thus, a bidirectional effect would be observed in both the L1 and L2 of our participant.

As mentioned, one of the variables under study was the modifications of MOA of rhotics in L1 and L2. In English, approximants are the most dominant allophones of /r/ (Laver, 1994; Lodge, 2009), whereas in Farsi different allophones of rhotics occur in different positions (Hall, 2007; Izadi, 2014;

Rafat, 2008; Samareh, 2002),). In English, stress in bi-syllabic nouns is often on the first syllable, whereas in Farsi stress in nouns is on the final syllable (Kreidler, 1987). Considering these features, it was predicted that;

1. the child's L1 will be attrited. Specifically, the MOA of phoneme /r/ and the correlates of SP in L1 Farsi will be assimilated to L2 English norms after spending more time in an English-speaking community. For example, [go.ˈɾaz] will be produced like [ˈgo.ɹaz] ‘wild boar’. So, after attrition, the tap /r/ will be an approximant. Moreover, before attrition the F0 peak is on the nucleus of the second syllable and the second syllable is longer, whereas, after attrition the peak will be placed on the nucleus of the first syllable and the first syllable will be longer than the second one.
2. The MOA of phoneme /r/ and the location of F0 and duration of the stressed syllable in L2 English will be produced according to Farsi norms, at the beginning. However, after being more in contact with English, the mentioned features will be produced in line with English norms. For example, at first *carrot* is produced as [kɛ.ˈɾət] with a tap /r/, while F0 peak is on the final syllable, and the final syllable is longer than the first syllable. But in the second session, /r/ will be produced more likely that of a native English speaker e.g., [ˈkɛ.ɹət].

Method

Participant

For the purpose of this study a 9-year-old female L1 standard Farsi user was examined. The participant’s age of arrival to Anglophone Canada was 9 years old, which is also, considered to be the onset of learning English as L2. She had received English language instructions in preschool, since she was 3 years old. She had also taken private lessons with a native English teacher for 2 years before coming to Canada.

Task

The experimental procedure included a picture-naming task and was carried out in two sessions that included Farsi and English. The first session was recorded 2 months after the arrival of the participant to Anglophone Canada. The second session took place one month after the first session.

Stimuli

In both sessions, the participant was presented with the same materials. 30 pictures were designed for the MOA of rhotics in different positions. Also, 20 pictures of bi-syllabic words were designed for the SP, 10 for Farsi and 10 for the English section. All the stimuli were repeated three times, generating a total of 300 tokens which were transferred to Praat (Boersma & Weenink, 2016) for the acoustic analysis.

Data analysis

The participant was audio-recorded as she was naming the pictures using an Audacity-macosx-ub-2.1.0 recorder. After the recording, a total of 300 tokens were extracted, labeled and transferred to Praat for acoustic analysis. Rhotics were coded as fricatives, taps, approximants, and deletions because /r/ was not produced in some positions. The length of all rhotics was measured as well. The acoustic analysis of the stimuli revealed that, in Farsi approximants and fricatives were produced word initially as in [ɹ̠oj] ‘lipstick’, fricatives post-consonantly as in [sefr̠] ‘zero’, fricatives, approximants and deletion of /r/ occurred in pre-consonant positions as in [faːʃ] ‘rug’ (in this example /r/ is deleted). Approximants and taps occurred intervocalically as in [suɹ̠æt] ‘face’, and approximants, fricatives and taps word finally, as in [ʔabʃaɾ] ‘waterfall’.

In English approximants and fricatives were released syllable initially, as in *rug* [ɹʌg], approximants and deletion occurred in onset cluster, as in *drum* [dɹʌm], approximants intervocalically, as in *carrot*

[ˈkɛ.ɹət], approximants and deletion in coda cluster, as in *fork* [foɹk], and approximants and fricatives, in syllable final position, as in *car* [kɑɹ]. After identifying the type of each rhotic the results of each session were compared with that of the second session to find out whether attrition has occurred or not, and also to examine whether the child made progress in acquiring English by producing approximant rhotics. The average length of rhotics in the first session and the second session was measured as well.

Regarding stress correlates, the results of measuring the F0 peak, and the duration of the syllables in the first session of the Farsi and the English tasks were compared to the results of the second session. Syllable length was determined by measuring the distance between the onset and offset of each syllable. The highest point of formant frequency on the nucleus of each syllable was considered as the F0 peak in each of the syllables.

Results

Rhotic production in Farsi and English

A total of 180 tokens were acoustically analyzed for distinguishing the type of rhotics: 90 tokens for the Farsi task and 90 tokens for the English task. The results of the acoustic analysis of Farsi rhotics yielded that 27% of rhotics were fricatives and 73% were approximants word initially. The same analysis in the second session showed that the production of approximants increased to 80%, while the production of fricatives decreased to 20%. Intervocalically in both sessions, 84% of rhotics were approximants and 16% were taps. In the pre-consonant coda position, 78% of rhotics were approximants. The other variants consisted of 11% fricatives and 11% /r/ deletion. In both sessions, only fricatives were distinguished in the final post-consonant position. In the first session, 89% of rhotics were fricatives and the other 11% were approximants in the final postvocalic position. However, in the second session the percentage of approximants increased to 45%.

For the English task, in both sessions, in the syllable initial position approximants and fricative were identified, being at 92% and 8% respectively. In onset cluster position, the percentage of approximants were 89% in both sessions and /r/ deletion was observed in the rest of the tokens. Intervocalically, all the rhotics were approximants in both sessions. In the syllable final post-vocalic position the results of the first session were: 74% approximants and 26% fricatives. In the second session, approximants increased to 86% and fricatives decreased to 7%, while 7% deletion of /r/ was also observed. In coda cluster position, 67% of rhotics were approximants and 33% were deleted in the first session. In the second session, approximants increased to 84% and deleted /r/ reduced to 16%.

An overview of the above results is given in the Figure 1. As can be seen in total, rhotics in the first session of the Farsi task consisted of 53% approximants, while in the second session this amount increased to 62%. Farsi fricatives were reduced from 43% in the first session to 27% in the second session. Also, the percentage of taps and deleted /r/ increased from to 7% and 4%, respectively. The increase in the number of approximants suggests that the child is producing English-like rhotics in her L1, which is evidence of L1 attrition.

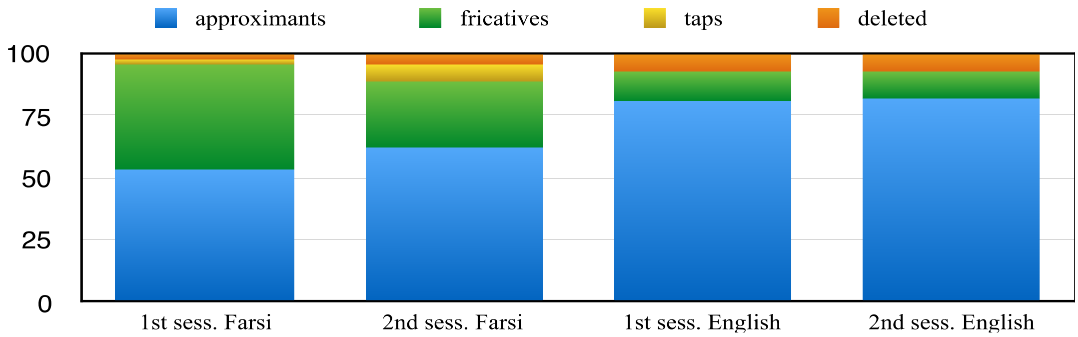


Figure 1. % of rhotics in Farsi and English of the bilingual child in all positions

Regarding the English task, almost no change can be seen in the manner of articulation of each type of rhotics from the first session to the second session. In both sessions, 80% of rhotics were approximants, revealing that the child has already showed progress in acquiring the production of English-like rhotics.

Stress production in Farsi and English

A total of 120 tokens, including 60 tokens for each language were acoustically analyzed for the purpose of measuring the correlates of stress pattern in L1 and L2 of the bilingual child.

Syllable duration in L1, Farsi

The syllable duration measurement in the first session illustrated that only in 7% of tokens the first syllable was longer than the second syllable. In the second session, this amount was almost doubled to 15%. As can be seen in Figure 2, in the first session the average length of first syllables were 272 ms and the average length of the second syllable was 462 ms. Whereas in the second sessions, the average length of the first and the second syllable was 277 ms and 315 ms, respectively. The decrease in the length of the second syllable of Farsi tokens in the second session suggests that the child's L1 is assimilating to her L2, which is evidence of L1 attrition.

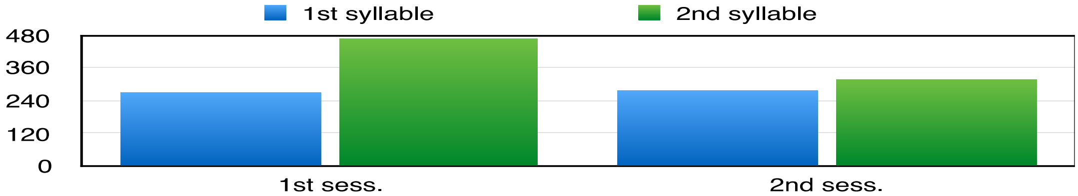


Figure 2. Syllable duration in MS in Farsi of the bilingual child

F0 peak in L1, Farsi

The results taken from the fundamental frequency measurements showed that in both sessions of the Farsi task, only in 3% of tokens F0 peak was on the nucleus of the first syllable, instead of being on the nucleus of the second syllable. Thus, no significant correlation was found between F0 peak and L1 attrition. As Figure 3 shows, in the first session the average high of F0 in the first syllables was 2746 Hz and the average high of the second syllable was 2520 Hz. In the second session, the average high of F0 in the first and second syllable was 3039 Hz and 3070 Hz, respectively.

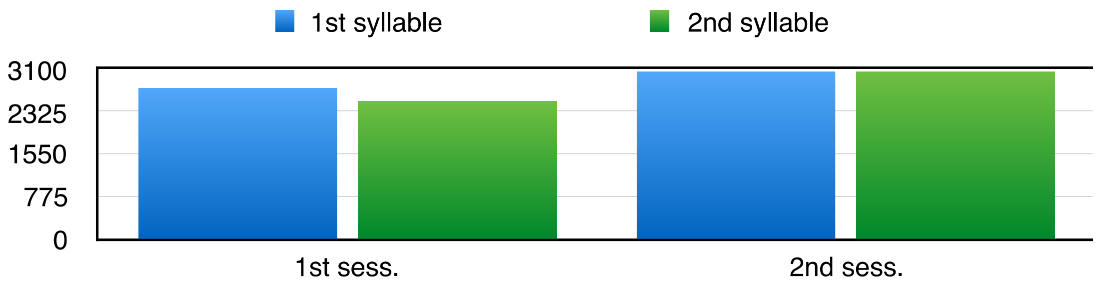


Figure 3. F0 peak frequency in Farsi of the bilingual child

Syllable duration in L2, English

The measurement of syllable duration indicates that in 20% of tokens the duration of first syllables was shorter than that of second syllables. This amount in the second session reduced to 7%. Figure 4 illustrates that in the first session the average length of first syllables was 349 ms and that of second syllables was 351 ms. In the second session, the average length of the first and the second syllable was 392 ms and 219 ms, respectively. The decrease in the length of second syllables in English tokens reflects that the bilingual child has produced more English like tokens. This is considered to be evidence of English acquisition.

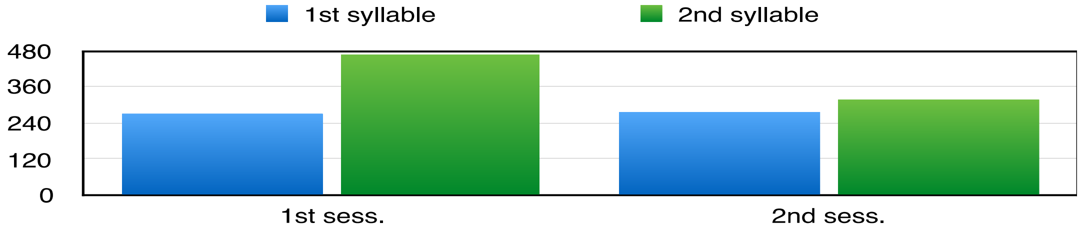


Figure 4. Syllable duration in Ms in English of the bilingual child

F0 peak in L2, English

The F0 peak measurement in English of the bilingual child demonstrates that in the first session only in 3% of the tokens the F0 peak height in the first syllable was longer than in the second syllable. Whereas, in the second session all F0 peaks were on the nucleus of the first syllable, which is in accordance with English SP norms. As can be seen in Figure 5, in the first session the average high of F0 in the first syllables was 2918 Hz and that of second syllables was 2369 Hz. In the second session, the average high of F0 in the first and the second syllable was 4546 Hz and 3299 Hz, respectively.

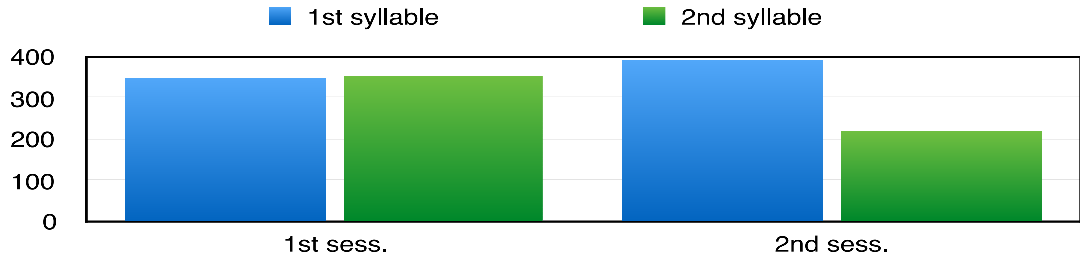


Figure 5. F0 peak frequency in Hz in English of the bilingual child

Discussion

In this study, it was hypothesized that L2 proficiency and L1 loss bidirectionaly affect each other (Major, 1992). This research examined this assumption by providing an acoustic analysis of rhotics and two correlates of stress pattern including F0 peak and duration of syllable in both the L1 and L2 of a bilingual child. It was predicted that in the first session of the Farsi task a number of different allophones of /r/ will be produced in different positions in the Farsi tokens. Yet, in the second session rhotics will mostly become approximants, suggesting that more English-like rhotics will be produced in L1. The reason is that approximants are the most frequent rhotics in Canadian English (McMahon, 2002). On the other hand, based on Kreidler's (1987) account of Farsi and English stress patterns, it was hypothesized that the position of stress in the second session will be assimilated to the English norms. Hence, the increase in the number of approximants in Farsi tokens in the second session will confirm that L1 attrition has taken place. Furthermore, as Major (1992), asserts that attrition and acquisition have mutual relation, if the number of approximants in English increases in the second session, and the mentioned correlates of stress alter from being in line with Farsi norms to be more like English norms after one month, then it can be confirmed that the child had made progress in acquiring English.

Results yielded that approximants and fricatives occurred word initially. Hence, as it was hypothesized that the increase in the number of the released approximants in the second session is a sign of L1 attrition, and the fact that an increase in approximants was observed, the first hypothesis, was confirmed in this position. No increase was detected in the rate of production of approximants intervocalically, thus, rhotics cannot give any clue in favor of or against the first hypothesis about L1 attrition. In the pre-consonant coda position, the occurrence of approximants is a typical feature of both Farsi and English. As a result, approximants occurred in both languages with a high rate of occurrence. Thus, rhotics in this position do not give any account of L1 attrition. The attrition hypothesis also was not confirmed in the post-consonantal position as none of the fricatives were changed to approximants. This could be due to the fact that the precedent consonants in the stimuli were fricatives and voiceless stops, and the production of fricatives right after them was easier for the participant due to similarity of place of articulation. As more approximants were released in final post-vocalic position in the second session, the first hypothesis on L1 attrition was confirmed.

Regarding stress correlates in Farsi: increasing the number of tokens in which the duration of the first syllable was more than the second syllable proved the first hypothesis. However, the results taken from the fundamental frequency measurements were not as consistent as syllable duration in admitting L1 attrition. So, in terms of supra-segmental level, duration is a more reliable factor to investigate the occurrence of attrition, since this correlate seems to undergo change faster than fundamental frequency.

In the English task, the number of approximants in syllable initial, onset cluster positions and intervocalically did not change. However, as the rate of their occurrence was high, it can be concluded that English rhotics in this position had already been acquired, so the second hypothesis was confirmed. In syllable final position and coda cluster, the increase in number of approximants confirmed the second hypothesis.

The result of syllable duration and F0 peak measurements revealed that, at the beginning, there was more production of English words with stress on their final syllable, whereas as time passed, this amount reduced as a result of being more in contact with English native speakers. Thus, the second hypothesis was confirmed.

Conclusion

This study acoustically analyzed rhotics and two correlates of stress including syllable duration and F0 peak in order to explore whether L1 attrition and L2 acquisition occur in the speech of a Farsi-English bilingual child or not. The total measurement of rhotics in Farsi revealed that in the second session the number of approximants increased in most positions. Thus, producing more English-like rhotics is considered to be evidence that phonological attrition is in progress in the speech of the bilingual child. Results of analyzing the duration of syllables in Farsi showed that more Farsi words were produced with a shorter second syllable, so the stress was altered to the first syllable. However, the F0 peak was not as consistent as duration in determining the changes in stress pattern in this study. Overall, considering all the mentioned criteria the first hypothesis was confirmed. The measurement of rhotics in English showed that in total the number of released approximants was the same in both sessions. Yet, the occurrence of this allophone of /r/ more than other types of rhotics is taken to be evidence that the MOA of rhotics in English was learned by the child. Moreover, the duration of syllables and location of F0 peak confirmed that the number of English words with correct stress patterns increased, so it can be concluded that the child has acquired English and thus the second hypothesis was confirmed too. Taken together, the results suggest that L2 acquisition fostered L1 attrition.

References

Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.22, retrieved 15 November 2016 from http://www.praat.org/.

Hall, M. (2007), *Phonological characteristics of Farsi speakers of English and L1 Australian English speakers' perceptions of proficiency*. Master's Thesis.

Izadi, M. (2014). Spanish and Persian's Taps and Trills: A Comparative Study. *Journal of Academic and Applied Studies*, *4*(2), 25-36.

Kreidler, C. W. (1987). Stress differentiation in sets of English words. *Word*, *38*(2), 99-125.

Laver, J. (1994). *Principles of phonetics*. Cambridge, UK: Cambridge University Press.
Lodge, K. (2009). *A critical introduction to phonetics*. London, UK: Continuum International.
Major, R. C. (1992). Losing English as a first language. *The Modern Language Journal*, *76*(2), 190-208.
McMahon, A. (2002). *An introduction to English phonology*. Oxford, UK: Oxford University Press.
Rafat, Y. (2008). A socio-phonetic investigation of rhotics in Farsi. *7th Biennial Conference on Iranian Studies*, University of Toronto, Toronto, ON.
Rafat, Y., Mohaghegh, M., & Stevenson, R. (2017), Geminate attrition across three generations of Farsi-English bilinguals living in Canada: An acoustic study. *Ilha do Desterro- Investigating Second Language Speech Learning*, *71*(3), 151-168.
Samareh, Y. (2002). *Ava shenasiye zabane Farsi*. Tehran: Makaze Nashre Daneshgahi.
Schmid, M. S. (2008). *Language attrition*. Cambridge, UK: Cambridge University Press.

To cite this article: Fakoornia, N. (2017). L1 Farsi attrition in contact with L2 Canadian English: A focus on rhotics and stress pattern. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 120-126). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Some acoustic characteristics of the geminates in the Ikema dialect of Miyako Ryukyuan

Masako Fujimoto[1], Shigeko Shinohara[2]
m.fujimoto5@kurenai.waseda.jp, shigeko.shinohara@sorbonne-nouvelle.fr

[1]Waseda University, [2]LPP, CNRS/Sorbonne-Nouvelle

Abstract. Ikema is an endangered Miyako Ryukyuan dialect spoken in the Okinawa islands of Japan, having a set of voiced/voiceless geminates in word-initial and medial positions. Among its obstruents, the geminates /tt ff ss vv zz tts/ appear in initial and medial positions (/tta/ 'tongue', /ffa/ 'child', /ssa/ 'grass', /vva/ 'you', /zzu/ 'fish', /avva/ 'oil') and the geminate /dd/ occurs only in medial position (/badda/ 'side'). This paper is an acoustic description of some aspects of these geminates, as a follow-up to the rt-MRI study by Fujimoto and Shinohara (2015), which showed articulatory differences between voiceless and voiced, as well as singleton and geminate consonants. In this study, we analyzed sets of real words uttered several times in isolation by five native speakers by using oscillograms and spectrograms. First, we confirmed longer closure duration for plosives and affricates and longer frication noise for fricatives in geminates as compared to singletons. Second, our acoustic analyses found full voicing of voiced geminates in Ikema. This result seems to be a consequence of pharyngeal expansion identified in the MRI study during voiced geminate constriction. Third, VOT was systematically shorter for the geminate voiceless plosive (/tt/) compared to its singleton counterpart. This corresponds to the faster transition from the geminate consonants to the following vowel found in the MRI study.

Keywords: geminates, Ikema Ryukyuan, pharyngeal expansion, VOT, mora-timing, dialect

Introduction

Ikema dialect is spoken in Miyako Island in southern Ryukyu, or Okinawa, Japan. UNESCO designated Miyako dialects as endangered. Ikema differs from Tokyo Japanese in that it has typologically rare voiced geminates as well as word-initial geminates. Table 1 presents geminate consonant types in Ikema Ryukyuan. It has voiceless and voiced geminates in word-initial and word-medial positions.

Table 1. Geminates in Ikema Ryukyuan

	Word initial	Word medial
Voiceless	*tta* 'tongue', *ffa* 'child', *ssa* 'grass', *ttɕutsI* 'cycada'	*uttu* 'brother', *attɕa* 'geta clogs', *ntts*-a 'street. TOP', *maffa* 'pillow', *uss*-a 'cow. TOP'
Voiced	*vva* 'you', *zza* 'father'	*badda* 'side', *avva* 'oil', *tuzz*-a 'wife. TOP'

We take Tokyo Japanese as a point of reference because it also has a distinction between geminate and singleton obstruents. /toppa/ 'breakout', /kitte/ 'stamp', /kissa/ 'tea drinking', /mattja/ [mattɕa] 'ground green tea' are some of the words with voiceless geminates. The phonetic difference between singletons and geminate obstruents in Tokyo Japanese is mainly due to duration of frication noise for fricatives and that of closure phase of plosives and affricates (Han, 1962 among others). Fujimoto and Shinohara (2013) report duration and other acoustics differences of word-initial obstruent singletons and geminates (/s/ vs. /ss/, /t/ vs. /tt/, and /z/ vs. /zz/) uttered in a frame sentence by a speaker of Ikema Ryukyuan. Frication noise was longer in geminate fricatives and closure duration was longer for the plosives. Voiced geminate obstruents appear only in recent loanwords in Tokyo Japanese and they are variably devoiced in realization (Kawahara, 2006; Matsuura, 2012). In contrast, according to Matsuura (2012) in a preliminary acoustic study with a single token by a single speaker, word-medial voiced geminate [vv] is fully voiced in Ikema Ryukyuan.

A preliminary analysis of real-time MRI (rt-MRI) data of Ikema Ryukyuan recorded in 2014 (Fujimoto & Shinohara, 2015) examined eight words uttered by a speaker. Among the findings reported in that study, the following aspects will be examined in the current acoustic studies. In rt-MRI, it was observed that linguo-palatal constriction was not only longer in duration but also larger in the area during geminates than singletons, as estimated from the dimension of the constriction of the tongue front against the hard palate for coronal obstruents. This result was similar to what was reported for coronal geminate obstruents in Tokyo Japanese (Kochetov & Kang, 2013) and in other languages (e.g., Ridouane, 2010). The longer and firmer constriction in articulation should result in longer frication noise for fricatives and longer closure duration for plosives and affricate geminates. Concerning the voicing of voiced geminate obstruent [z], pharyngeal expansion was observed during the articulation of /zz/ but not during voiceless /ss/. An expanded pharynx facilitates vocal-fold vibration (Perkel, 1969 among others). Thus, we expect that there will be clear acoustic manifestation of voicing during voiced obstruents of Ikema Ryukyuan. Another observation was the faster transition from coronal obstruents to a following vowel in geminates than in singletons, such as in /tta/ [tta] than that in /ta/ [ta:]. This may cause the acoustic outputs such as VOT to differ between geminate and singleton.

We shall investigate the acoustic characteristics of geminates based on the observations in the rt-MRI in Ikema Ryukyuan. First, we observe whether geminate consonants show longer acoustic duration than their singleton counterparts as in Tokyo Japanese. Second, given that geminates have been shown to have longer duration than singletons, we investigate whether Ikema shows timing control proportional to mora count as in Tokyo Japanese rather than to syllable count (Kondo & Shinohara, 2003; Sagisaka & Tohkura, 1984). Third, in order to find out acoustic voicing patterns, we observe voiced obstruents on spectrograms and oscillograms. Finally, we measure VOT of voiceless singleton and geminate stops to see whether VOT measurements reflect faster tongue transition after a geminate consonant as observed in the rt-MRI study.

Ikema dialect

The consonant inventory of Ikema dialect is /p b t d k g ts s z f v h m n r j w N N̥/ (adapted from Pellard & Hayashi, 2012). These phonemic symbols roughly reflect phonetic values in IPA unless mentioned otherwise. /N/ is a placeless moraic nasal. /N̥/ realizes as a nasal fricative that assimilates its place to the following consonant. /r/ is typically an alveolar tap but many variants also exist. [ts] and [tɕ] are in complementary distribution in Ikema Ryukyuan. And /s/ and /z/ are palatalized before /i/ and /j/ ([ɕ] and [ʑ], respectively). Among these consonants, single obstruents occur with [p b t d k g ts tɕ s z ɕ ʑ f h] in word-initial and word-medial positions. Geminate obstruents occur with [ff vv tt tts ttɕ dd ss ɕɕ zz ʑʑ] in initial and/or medial positions. Note that no single [v] occurs anywhere and geminate [dd] occurs only in word-medial position. The vowel inventory is /a i u I/ where the phone /I/ denotes a central/apical vowel possibly with frication noise. The vowel /o/ appears in a very limited number of words. There is a length distinction for vowels, except in monosyllabic words with a singleton consonant which are pronounced with a lengthened vowel by a bimoraic minimal word constraint. That is, a word must have at least two moras so that underlingly monomoraic words surface as having two moras with a lengthened vowel: /ti/ [ti:] *[ti] ‘hand’. Thus, monosyllabic words with an initial geminate consonant are not pronounced with lengthened vowel (cf. /tta/ [tta] ‘tongue’). Geminate consonants are assumed to be moraic.

Method

We recorded one female and four male speakers of Ikema Ryukyuan. They were in their 60’s and 70’s at the time of recording. All participants were bilingual between Standard Japanese and Ikema Ryukyuan. We selected around 100 lexical items based on a previous study by Kibe (2012) for several aspects of sounds of this dialect including the current one on geminates. Test words in Ikema dialect and their translations in Standard Japanese were prepared as a written list. The speakers were asked to translate each word presented orally in Standard Japanese by the experimenter into Ikema dialect. Three of the participants were asked to pronounce the word three times in isolation and three times in a frame sentence, *uryaa* __ *do* /urjaa__do/ 'I say __'. The other two older participants pronounced each word once in the following format: ___, *uryaa* ___ *do*. They were asked to keep the tempo as constant as possible and not to make a pause within the sentence during the production. In the latter format, a short pause was inserted between the word (in isolation) and the frame sentence. The difference in format was due to availability of the speakers. Speakers at times gave some of the items with different forms from the expected ones, which resulted in some variability of test words among participants. Utterances were recorded using a Sony PCM-D50 recorder with integrated microphones at a sampling rate of 48Hz, 16bits. Recording took place at the speaker's house in the Hirara quarter of Miyako Island in 2011 for four speakers and at a recording room in the University of Kyoto in 2014 for one male speaker. One of the speakers was a participant of the rt-MRI study in 2014. Among the collected data, we used relevant word sets suitable for examining duration and voicing. Since we used only real words, word pairs were constructed as minimally different as possible; however, not all possible combinations were available. In such cases, near-minimal sets were used. We analysed words in isolation using Praat (Boersma & Weenink, 2016). The number of tokens per word mentioned in the following sections are eleven (three tokens each by three speakers and one token each by two speakers) with variation due to some extra or failure tokens, unless otherwise noted. The word sets will be introduced in relevant sections. The rest of the items are not studied in this paper.

Results

Duration pattern

Because we could not find minimal pairs of geminate and singleton fricatives, we made comparisons between the geminate /ss/ in /ssa/ [ssa] 'grass' and the singleton /s/ in /sata/ [sata] 'sugar', and also between /ff/ in /ffa/ [ffa] 'child' and /f/ in /fau/ [fau] 'to eat' (or in /fai/ [fai], a variant, by one speaker). All are two mora words. According to duration measurements from spectrograms (see Figure 1), frication is longer in the geminate with fricative consonants. The averaged durations with SD in parentheses followed by the number of consonants are as follows: For /f/, 90 ms (17ms), N=8; for /ff/, 178 ms (14ms), N=11; for /s/, 176ms (17ms), N=8; and for /ss/, 211ms (14ms), N=12. The result of ANOVA with consonant (either /s/ or /f/) and gemination (either singleton or geminate) as independent variables and consonant duration as dependent variable showed that consonant ($F(1,35)=15.362$, $p=0.000$) and gemination ($F(1,35)=16.296$, $p=0.000$) were both significant but interaction was non-significant ($F(1,35)=2.962$, $p=0.094$). Thus, fricative geminates are significantly longer than singletons.

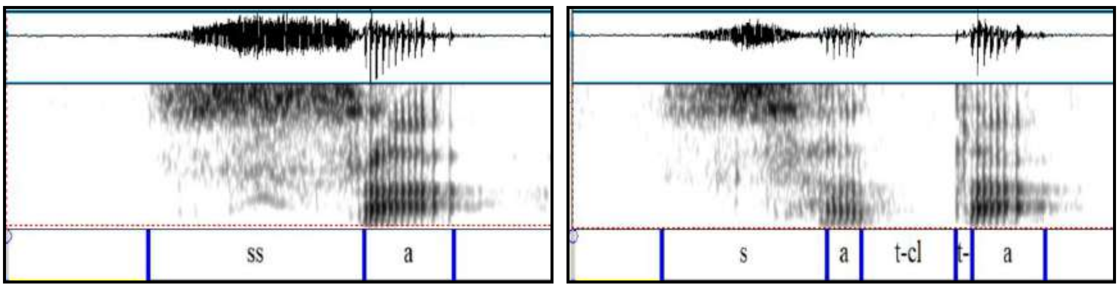


Figure 1. Spectrograms of /ssa/ [ssa] 'grass' vs. /sata/ [sata] 'sugar'. Frame length is 700 ms

We also observed a qualitative difference in sharpness of the boundary between the frication noise and the onset of the first vowel. As can be seen in Figure 1, the boundary is sharper after the geminate than after the singleton. We also observed that the vowel /a/ following the fricative in /sata/ was devoiced in three tokens by two speakers, but none in /ssa/. This might suggest vowel quality differences due to the gemination of the preceding consonant. Further analysis is necessary to examine this point.

Plosives and affricates are also longer in geminates especially in the closure period. We compared the averaged closure durations of /ts/ [tɕ] in /atsa/ [atɕa] 'tomorrow' and /tts/ [ttɕ] in /attsa/ [attɕa] 'geta clogs', and /d/ [d] in /nada/ [nada] 'tear' and /dd/ [dd] in /badda/ [badda] 'side'. The average duration of the closure period, SD in parentheses and the number of tokens of the consonants were 63 ms (SD 19 ms) N=12 for /ts/, 180 ms (SD 44 ms) N=11 for /tts/, 49 ms (SD 8 ms) N=11 for /d/, 171 ms (SD 28 ms) N=11 for /dd/. Results of ANOVA with consonants (either /(d(d)/ or /(t)ts/) and gemination (either singleton or geminate) as independent variables and the closure duration as dependent variable showed that closure duration was significantly longer in geminates than in singletons ($F(1, 41)=208.859$, p=0.000), but consonants and interaction was non-significant. Thus, we find acoustic duration differences in geminate segments from singleton counterparts in these plosive and affricate pairs.

In order to examine moraic timing in Ikema Ryukyuan, we measured the word duration of the test words of two speakers, classifying a geminate consonant or a light syllable as one mora, a combination of a geminate consonant and a short vowel as a heavy syllable (two moras), and a long vowel or a vowel followed by a placeless nasal N also as a heavy syllable (2 moras). We present here the detailed results of one of the speakers. Table 2 shows the classification of the test words (those shorter than four moras) into four groups, according to their syllable/mora breakdown. Word durations were 441 ms (SD 25 ms), N=5 for 3 syllable/3 mora words ((C)VCVCV), 350ms (SD 12 ms), N=21 for 1 syllable/2 mora words (CCV/CVV/CVN), 454 ms (SD 10 ms), N=30 for 2 syllable/3 mora words (CCVCV/(C)VCCV/VCVN) and 390 ms (SD 14 ms), N=15 for 2 syllable/2 mora words ((C)VCV). Figure 2 shows average durations of words uttered in isolation by a speaker. ANOVA with mora count and syllable count as independent variables and word duration as a dependent variable showed that mora count was a significant factor for word duration ($F(1, 67) = 13.231$, p=0.000), but not syllable count was ($F(1, 67) = 2.314$, p=0.107). (Interaction of mora and syllable is not applicable as it has a zero degree of freedom). The other speaker also showed a significant difference between 3 mora groups and 2 mora groups of words with a partly different word list. The details are omitted due to space limitations. The results indicate moraic timing in this dialect.

Table 2. Ikema words differing in syllable and mora counts for duration comparison

3 syllables/3 moras	*isagu* 'cough', *fuzata* 'brown sugar'
2 syllables/3 moras	*akaN* 'dirt', *maffa* 'pillow', *NnutsI* 'life', *nuuma* 'horse', *attsa* 'geta', *adaN* 'a type of plant', *ffatsI* 'hoe', *zzaku* 'oar', *NNta* 'nut'
2 syllables/2 moras	*munu* 'thing', *naba* 'dirt', *nada* 'tear', *sata* 'sugar', *atsa* 'tommorow'
1 syllables/2 moras	*miN* 'ear', *Nna* 'rope', *ssa* 'grass', *N̥mu* 'cloud', *Nta* 'soil', *NtsI* 'road', *zI* [zI:] 'soil'

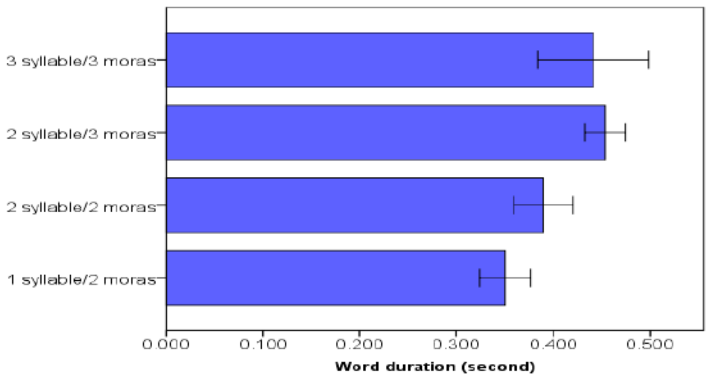


Figure 2. Three mora- vs two mora- word durations uttered in isolation by a speaker. Error bars show the 95 % confidence interval

Voicing pattern

Regarding voicing patterns, two characteristics were observed. The first one is the full voicing of voiced geminate obstruents. Voiced geminates presented voice bars during the frication for fricatives and during closure and release for plosive /d/ for 75 tokens (seven words * three tokens * three speakers + six words * one token * two speakers). In addition, a long pre-voicing often preceded initial voiced geminate consonants, as illustrated in the left-side of the spectrogram of [v] in /vva/ [vva] 'you' in figure 3. Pre-voicing occurred seven tokens out of 11 among five speakers, duration of which was 81 ms on average (SD 37 ms). For word-medial voiced plosives, voicing continued through the whole closure phase as illustrated in the right-side spectrogram of /badda/ [badda] 'side'. This is true for most of the 11 tokens with an exception of a female speaker where [dd] was partially devoiced before release.

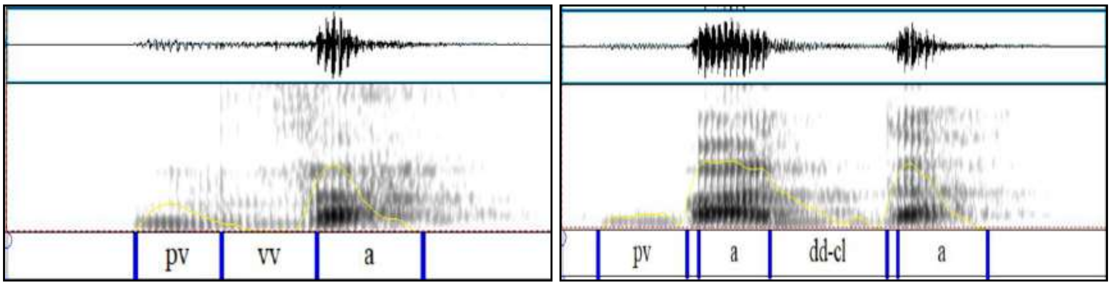


Figure 3. Spectrograms of /vva/ [vva] 'you' (left), /badda/ [badda] 'side' (right). Frame length is 700 ms

The second characteristic is the VOT difference between geminate and singleton voiceless stops. A voiceless geminate plosive has a shorter lag. Singleton stops have noticeably long VOT. We examined the VOT differences of [t] between /tta/ [tta] and /ta/ [ta:] and [t] between /ttii/ [tti:] and /ti/ [ti:] 'hand'. Note that the vowels in /ta/ and /ti/ are lengthened by the bimoraic minimality constraint. Each word was uttered in isolation by five speakers (11 tokens for each word: three tokens by three speakers and one token by two speakers). VOT is on average 60.4 ms (SD 45.3 ms) for /ta/, and 0.6 ms (SD 4.8 ms) for /tta/, 48.7ms (SD 11.3 ms) for /ti/ and 16.4 ms (SD 4.6 ms) for /ttii/. A t-test showed that the difference is significant between /ta/ and /tta/ ((t (20) =3.626, p=0.002), and between /ti/ and /ttii/ (t(19)=8.443, p=0.000).

Discussion

We have presented some acoustic characteristics of geminate obstruents of the Ikema dialect of Miyako Ryukyuan. Our main findings concern duration and voicing patterns.

In the phonological description of Ikema Ryukyuan, the geminate consonants are considered to be moraic (Pellard & Hayashi, 2012). However, phonetic realizations of geminate consonants in this dialect have not been fully understood. Our previous study with rt-MRI of geminate obstruents in Ikema Ryukyuan made the following observations: 1) longer and firmer constrictions for geminates, 2) pharynx expansion for voiced geminate obstruents and 3) faster movement of articulators from a voiceless geminate onset to the following vowel. The current study examined the recordings of five native speakers to identify the acoustic correlates of the previously reported articulatory characteristics, and also, to examine related issues. We compared durations of geminate consonants with those of singletons with several (near)-minimal pair words including those involving initial geminates. In addition, we also investigated whether the mora or the syllable is the unit of timing control at the word level in this dialect. Another phonological characteristic of this dialect is the existence of voiced geminate obstruents. Voicing through a voiced geminate [vv] was observed in the acoustic signal of a previous study (Matsuura, 2012). The current study also observed voicing during frication of two types of fricatives, and during the closure phase of voiced plosive /dd/. Another aspect related to voicing is the VOT difference of voiceless plosives /tt/ and /t/. Since the rt-MRI study indicated a faster articulator movement after the geminate /tt/, we expected some kind of acoustic manifestation accompanying it. The present study found a shorter lag for /tt/, clearly different from the lag for the singleton /t/ in two minimal word pairs.

Duration differences were observed for a small set of consonants (/s/ vs. /ss/, /t/ vs. /tt/, and /z/ vs. /zz/ in a frame sentence) in Fujimoto and Shinohara (2013). The current study extended the findings of geminates being acoustically longer than singleton counterparts to more various types of consonants. More precisely, geminate fricatives [ss] and [ff] in initial position showed longer frication noise. Geminate plosive [dd] and affricate [ttɕ] in word-medial position showed longer closure duration. These results corroborate with previous studies showing longer frication noise in geminates [ss] and [zz] and longer closure duration in the voiceless geminate stop [tt]. The results on obstruent duration indicate that phonological contrast between geminate and singleton consonants are realized by phonetic duration differences in Ikema Ryukyuan, as in Tokyo Japanese.

A question arises as to whether Ikema Ryukyuan is a mora-timed language similar to Tokyo Japanese. We compared duration of words classified with numbers of syllables on the one hand and with the number of moras on the other. The words differing in number of syllables behaved similarly when their number of moras was the same. For instance, CVCCV (two syllable/three mora) word (e.g., /maffa/ ‘pillow’) showed a more similar duration to CVCVCV (three syllable/three mora) words (e.g., /fuzata/ ‘brown sugar’) than to CVCV (two syllable/two mora) words (e.g., /sata/ ‘sugar’). The result of the timing investigation indicated that a speaker’s utterance is based on mora rhythm, at least at the word level, since mora numbers had a significant effect on the word duration whereas syllable numbers did not always. Speakers uttered their test words incited by oral translation; thus, there is no reason to consider any influence from the writing system. Moreover, Ikema dialect is largely an oral language in which there is not even a standard way of using Japanese syllabary to write in the dialect. Since Ikema Ryukyuan seems to operate on moraic timing, longer duration of geminate consonants must reflect its moraicity.

The voicing pattern, on the other hand, differs from Tokyo Japanese. Voiced geminate obstruents are used only in a marginal vocabulary in Tokyo Japanese and they are reported as being devoiced. Avoiding voiced geminate obstruents seems to be a cross-linguistic tendency (Jaeger, 1978; Kirchner, 2000; Podesva, 2002). Although voicing throughout geminate obstruents is physiologically challenging (Rothenberg, 1969; Westbury, 1983), articulatory settings enabling such production are partly revealed in rt-MRI studies. Speakers expand the pharynx during voiced geminates presumably to initiate and to keep the voicing (Fujimoto & Shinohara, 2015, 2017). The current study found that initial voiced fricative geminates [vv] and [zz] are fully voiced with pre-voicing, and voiced geminate stop [dd] is voiced during the closure (with only one exception out of 11 tokens).

Voiceless stops in Ikema Ryukyuan are different between singletons and geminates in terms of VOT. We examined /t/ vs. /tt/ in initial position with two minimal word pairs. Voiceless singleton plosives in both pairs had longer aspiration than geminate counterparts did. It was found in the rt-MRI study that geminate plosives had faster transitions toward the following vowel target (Fujimoto & Shinohara, 2015). One of the acoustic correlates of this articulatory difference must be the VOT. VOT might signal whether the voiceless stop is a singleton or a geminate, since the silent closure duration is useless as a cue in word-initial position (note that Ikema dialect does not have length contrast of the labial stop, which could have been cued visually). The Ikema dialect contrasts voiceless alveolar singleton /t/ and geminate /tt/ but lacks the initial voiced geminate /dd/. Lack of /dd/ in initial position can be explained by its phonetic similarity with /tt/. Since VOT is short after /tt/, it would be easily perceptually confused with /dd/ if they both occurred in initial position. However, in word-medial position other cues such as closure duration differences before voiceless and voiced stops might be used for perception, so that the contrast can be maintained.

The rich inventory of voiced geminates in the lexicon requires clear contrasts between voiced and voiceless counterparts. Voiced and voiceless fricatives may be differentiated by full voicing of the former. The same is true for word-medial alveolar plosives. Initial geminate /tt/ had a very short lag. Initial voiceless singleton and geminate plosives may be distinguished by their VOT difference.

Future directions may include articulatory, acoustic and perception studies. A more detailed study of the articulatory correlates of voiced geminates is in progress. Further acoustic investigation may include analyses of the burst spectrum of obstruents and formants of the following vowel as they might also cue voiceless initial geminates in addition to VOT differences.

Acknowledgements

We thank an anonymous reviewer and Donna Erickson for useful comments. This study was supported by SOLIFIC, Sophia University, JSPS Kakenhi No. 26370470 and 17K02707 and LABEX EFL.

References

Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.22, retrieved 15 November 2016 from http://www.praat.org/.

Fujimoto, M., & Shinohara, S. (2013). Acoustic characteristics of geminate consonants in the Ikema variety of Miyakojima dialect of Southern Ryukuan. *Proceedings of Autumn meeting of Acoustical Society of Japan*, 431-432. (in Japanese)

Fujimoto, M., & Shinohara, S. (2015). Articulation of voiced geminate and word-initial geminate obstruents in the Miyakojima Ikema dialect of Ryukyuan: A real-time MRI analysis, *GemCon 2015* Glasgow, 2015, poster presentation.

Fujimoto, M., & Shinohara, S. (2017). Articulatory manifestation for voiced geminates and initial geminates in Ikema Ryukuan: Evidences of rt-MRI data. *Methods in dialectology XVI*, Tachikawa, Japan.

Han, M. (1962). The feature of duration of Japanese. *Onsei no Kenkyu* [Study of sounds], *10*, 65-75.

Jaeger, J. (1978). Speech aerodynamics and phonological universals. In *BLS 4*, 311-329.

Kawahara, S., (2006). A faithfulness ranking projected from a perceptibility scale: The case of [+voice] in Japanese. *Language*, *82*, 536-574.

Kibe, N. (ed.) (2012). General study for research and conservation of endangered dialects in Japan: Research report on Miyako Ryukyuan. *NINJAL Collaborative Research Project Reports*, 12-02, 2012. (in Japanese)

Kirchner, R. (2000). Geminate inalterability and lenition. *Language*, *76*(3), 509-545.

Kochetov, A., & Kang, Y. (2013). When long is tense and tense is long: An EPG study of Japanese and Korean stops. Paper presented at *International Conference on Phonetics and Phonology 2013*, Tachikawa, Japan..

Kondo, M., & Shinohara, S. (2003). Durational evidence of psychological reality of the mora in Japanese speakers' French. *Proceedings of the 15th International Congress of Phonetic Sciences*, 463-466.

Matsuura, T. (2012). Miyako shohogen no onseijitugen ni kansuru yobiteki kento [Preliminary study on phonetic realization in Miyako dialects. In Y. Kibe (ed.), *General study for research and conservation of endangered dialects in Japan - Research report on Miyako Ryukyuan* (pp. 111-126).

Pellard, T., & Hayashi, Y. (2012). Miyako shohogen no onin- Taikei to hikaku- [Phonology of Miyako diaects – System and comparison]. In Y. Kibe (ed.), *General study for research and conservation of endangered dialects in Japan - Research report on Miyako Ryukyuan* (pp. 13-52).

Perkel, J. S. (1969). *Physiology of speech production: Results and implications of a quantitative cineradiographic study*. Cambridge, MA: MIT Press.

Podesva, R. (2002). Segmental constraints on geminates and their implication for typology. Talk presented at the *76th Annual Meeting of the Linguistic Society of America*, San Francisco.

Ridouane, R. (2010). Gemination at the junction of phonetics and phonology, Paper presented at *Labphon X*.

Rothenberg, M. (1969). The breath-stream dynamics of simple-released-plosive production. *Bibliotheca Phonetica 6*. Basel: S. Karger.

Sagisaka, Y., & Tohkura, Y. (1984). Phoneme duration control for speech synthesis by rule (in Japanese). *Journal of the Electronic Information and Communication Engineering*, *67-A*, 629-636.

Westbury, J. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing, *JASA*, *73*(4), 1322-1336.

To cite this article: Fujimoto, M., & Shinohara, S. (2017). Some acoustic characteristics of the geminates in the Ikema dialect of Miyako Ryukyuan. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 127-133). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Morphologically relevant speech processing skills in Greek-speaking children

Eleftheria Geronikou[1], Bill Wells[2]
e.geronikou@teiwest.gr, bill.wells@sheffield.ac.uk

[1]T.E.I. of Western Greece, [2]University of Sheffield

Abstract: While it is known that language development is related to auditory perception (Ziegler, Pech-Georgel, George, Alario, & Lorenzi, 2005) and that language difficulties frequently co-occur with speech difficulties (Broomfield & Dodd, 2004), there is currently little knowledge about the development of morphology in relation to the speech processing system. In this study a psycholinguistic framework for speech processing (Stackhouse & Wells, 1997) was adopted to investigate the development of phonological and morphological skills in children acquiring Greek. It was investigated whether morphological items pose specific challenges in terms of speech processing. Two groups of typically developing children aged 3;0-3;5 and 4;6-5;0 years respectively were assessed longitudinally at three assessment points six months apart. A range of a) phonologically related speech processing tasks and b) morphologically related speech processing tasks were used. Morphological phenomena e.g., tense and possessive pronouns were chosen as the basis for experimental stimuli of minimal morphological difference. These stimuli were used to assess input and output processing in tasks of speech perception and word and nonword repetition with matched items across tasks, so that comparisons could be made. On most of the matched tasks there was no significant difference between morphological and phonological results. Moreover, a significant relationship was found between processing of phonological and morphological items. Assessment of two children aged 4;2 and 5;7 who had difficulty with the production of /s/, a phoneme used in multiple phonological and morphological contexts in Greek, revealed ≤ -1.5 S.D. from age matched controls performance on several tasks. Cross-domain generalization was observed upon intervention. In conclusion, based on longitudinal and single case studies data, it is suggested that morphological items, compared to phonological items, do not pose specific challenges in terms of speech processing, indicating that morphological characteristics of spoken language are an integral part of lexical representations. Morphology thus merits attention in assessment and intervention for speech sound disorders.

Keywords: speech processing, phonology, morphology, Greek

Introduction

There is a growing body of literature demonstrating that speech disorders and language difficulties frequently co-occur in children (Broomfield & Dodd, 2004; Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997). Epidemiological studies indicate that speech difficulty was a co-occurring issue for two-thirds of children with language comprehension difficulties and over half of children with expressive language difficulties (Broomfield & Dodd, 2004). Speech perception, such as consonant identification has been found to relate to the comprehension and production of spoken language (e.g., Ziegler *et al*. 2005). However, there is currently little knowledge about the development of morphology in relation to the speech processing system. This was addressed in the present study by adopting a psycholinguistic framework for speech processing (speech input, speech output, and lexical representations, as described by Stackhouse and Wells (1997) in order to assess the development of speech processing skills in typically developing children and in children with speech disorders who acquire Greek.

Greek is a highly inflected language. A variety of morphemes are used to indicate gender, number and case for nouns, person, number, tense and voice for verbs (Holton, Mackridge, & Philippaki-Warburton, 1997). Investigating speech processing and the development of morphology in a morphologically rich language, such as Greek, may inform theories of language acquisition, speech processing and clinical practice.

There is a growing body of literature suggesting that the principles of psycholinguistics can be applied in research of typical (Vance, Stackhouse, & Wells, 2005) and atypical (Constable, Stackhouse, & Wells, 1997; Rees, 2009; Vance, 1997) speech development. Most of the existing literature has studied English speaking children. However it is important to use this approach to study languages other than English in order to formulate and test hypotheses about the speech processing system cross linguistically. Within the simple psycholinguistic paradigm, it is supposed that a number of input processes occur when a child listens to spoken language and a number of speech output processes occur when a child speaks. Linguistic information is also stored in the form of phonological, semantic and grammatical representations that enable a child to understand and produce spoken language. The psycholinguistic perspective is adopted to investigate the development of phonological and morphological skills, particularly whether morphological items pose specific challenges in terms of speech processing.

Morphemes certainly carry grammatical information; each morpheme consists of one or more particular sounds that differentiate it from the others. It is explored whether the developmental pattern for the acquisition of morphological and phonological elements is similar or significantly different and whether performance in speech processing tasks is related in morphological and phonological items. The central hypothesis of the study is that the successful acquisition of the phonological i.e., perceptually distinct units of sound that differentiate the meaning of words and the morphological characteristics i.e., meaningful, grammatical units of spoken language, depends on the accuracy and efficiency of speech processing skills.

Normative Study

Method

Design

A cross sectional longitudinal study investigating typical speech development in two groups of Greek children aged three to six years old was undertaken. Children were assessed at three time points, six months apart, so that any change observed could be attributed to development.

Participants

Two groups of typically developing children participated in the study. Group 1 (16 participants) aged 3;0-3;5 at the beginning of the study were children attending a day care setting and Group 2 (22 participants) aged 4;6-5;0 were children attending a kindergarten school. All children had Greek as their primary language. Bilingual children were included under the condition that one of the parents was a native Greek speaker and that the child had been attending a Greek-speaking school setting for at least six months. All children passed a hearing screening test. Participants either had no vision problems or had vision problems that were corrected with glasses. In order to ensure that there were no structural or functional abnormalities of articulators a diadochokinetic task was used.

Tasks and materials used

The evaluation material included (i) published language assessments and (ii) experimental tasks of speech processing, where items of phonological and morphological interest were included with matched items across input and output tasks so that direct comparisons of performance could be done.

Published language tests

The Diagnostic Verbal IQ test (DVIQ, Stavrakaki & Tsimpli, 2000) was used as a baseline assessment of language development, to ensure that children had typically developing language skills. Three subtests were used: (i) the production of morphology and syntax (DVIQP), (ii) comprehension of morphology and syntax (DVIQC) and (iii) sentence repetition (DVIQSR).

Experimental tasks

The following experimental tasks of speech processing were used:

1. Nonword auditory discrimination, to assess the discrimination of speech sounds without reference to (a) phonological (NWAudDP) and/ or (b) morphological (NWAudDM) representations.
2. Real word auditory discrimination, to assess the ability to discriminate between words with different (a) phonological (RWAudDP) and (b) morphological (RWAudDM) elements from auditory presentation only.
3. Real word repetition, to assess the ability to produce (a) phonological (RWRepP) and (b) morphological (RWRepM) elements, when a model is given and stored linguistic knowledge may be used to support performance.
4. Nonword repetition, to assess the ability to produce sounds related to (a) phonological (NWRepP) and (b) morphological (NWRepM) elements without reference to representations.
   Performance on repetition tasks was scored for whole word (WW) and percentage of consonants correct (PCC) accuracy.

Experimental stimuli: Phonological minimal pairs

Phonological minimal pairs as [kuˈbi] ‘button’ [kuˈpi] ‘paddle’ were used to evaluate processing of perceptually distinct units of sound that distinguish one word from another in Greek. Voicing, manner and place of articulation and the phonotactic structure such as consonant clusters or close syllables were taken into consideration, in order to ensure broad representation of the Greek phonological system. Matching nonwords, like [kɐˈbi], [kɐˈpi], were created.

Experimental stimuli: Morphological minimal pairs

Morphological minimal pairs were used to evaluate processing of meaningful units of language that change the grammatical function of a word such as present/future tense, masculine/feminine gender. Minimal phonological difference may be found in morphological minimal pairs as [tɐˈizun] ‘are feeding’, [tɐˈisun] ‘will feed’. Matching nonwords were created by keeping the real word component that manifests the morphological difference intact and changing the vowel in the word stem [tɛˈizun], [tɛˈisun].

Table 1. Performance accuracy for each age group on published and experimental tasks

Task	Maximum Score	Group 1			Group 2		
		T1	T2	T3	T1	T2	T3
		3;0-3;5	3;6-4;0	4;0-4;5	4;6-5;0	5;0-5;5	5;6-6;0
		M (S.D)	M (S.D)	M (S.D)	M (S.D)	M (S.D)	M (S.D)
DVIQ P	24	8.19 (3.82)	10.81 (4.34)	15.11 (3.01)	11.31 (3.13)	15.91 (4.18)	16.16 (2.29)
DVIQ C	31	14.95 (3.15)	18.93 (3.33)	21.02 (3.61)	19.52 (5.02)	24.82 (2.89)	26.02 (2.23)
DVIQ SR	48	29.31 (12.51)	42.52 (4.81)	45.75 (2.32)	43.21 (5.41)	45.93 (2.71)	45.21 (5.09)
RWAudDP	30	19.06 (4.86)	21.81 (4.15)	24.86 (2.47)	23.90 (3.99)	25.47 (1.90)	27.39 (1.94)
RWAudDM	30	19.18 (3.70)	22.06 (3.73)	26.31 (2.91)	24.72 (3.99)	26.31 (3.63)	27.44 (2.87)
NWAudDP	30	18.56 (3.98)	19.18 (2.66)	23.50 (3.27)	22.54 (3.75)	25.84 (2.95)	26.05 (2.75)
NWAudDM	30	19.81 (3.45)	21.63 (3.81)	25.88 (3.07)	25.55 (3.75)	27.57 (2.11)	28.00 (2.00)
RWRepPWW	30	17.38 (7.77)	20.69 (7.11)	24.60 (5.95)	22.50 (6.49)	26.79 (5.11)	27.00 (3.66)
RWRepMWW	30	16.50 (7.64)	20.88 (9.07)	22.93 (8.19)	22.73 (5.73)	25.89 (7.32)	26.11 (5.95)
RWRepPPCC	100%	79.63 (16.52)	86.61 (12.85)	92.47 (9.21)	92.40 (8.17)	95.83 (6.84)	96.10 (4.76)
RWRepMPCC	100%	84.40 (15.20)	90.79 (10.51)	93.23 (8.57)	96.04 (7.65)	96.02 (7.82)	96.48 (5.60)
NWRepPWW	30	16.19 (7.54)	19.75 (7.13)	22.53 (5.94)	23.80 (5.46)	25.30 (5.19)	25.67 (4.46)
NWRepMWW	30	14.75 (7.34)	16.69 (7.74)	21.73 (7.96)	23.47 (7.97)	24.00 (7.12)	24.37 (6.44)
NWRepPPCC	100%	77.68 (15.84)	84.17 (15.08)	89.44 (9.12)	90.94 (8.48)	94.05 (6.93)	95.08 (4.85)
NWRepMPCC	100%	77.23 (17.30)	86.57 (12.56)	89.44 (11.47)	92.29 (11.31)	92.92 (10.62)	92.85 (8.61)

Results

Means and standard deviations were calculated for number of items correct for each age group at each assessment point (Table 1). Kolmogorov Smirnov and visual inspection of the histograms were used to ensure that data were normally distributed and Repeated Measures Anova was used to compare performance. Pairwise comparisons (with Bonferroni adjustment for multiple comparisons) were used to compare performance between time points.

Nonword auditory discrimination

In nonword auditory discrimination analysis yield a main effect of time for items of phonological ($F_{(2, 14)}$ =9.86, p= .002) and morphological ($F_{(2, 14)}$ =15.23, p<.001) interest for Group 1. Pairwise comparisons yield a significant difference between T1 and T3 (p=.003), T2 and T3 (p=.008) for items of phonological interest and a significant difference between T1 and T3 (p<.001), T2 and T3 (p=003) for items of morphological interest. For Group 2, there was a main effect of time for items of phonological ($F_{(2, 16)}$ = 15.25, p<.001) and items of morphological ($F_{(2, 16)}$ = 5.07, p=.020) interest. Pairwise comparisons yield a significant difference between T1 and T2 (p= .005), T1 and T3 (p < .001) for items of phonological interest and between T1 and T3 (p= .016) for items of morphological interest.

Real word auditory discrimination

In real word auditory discrimination, analyses yield a main effect of time for words with different phonological ($F_{(2,13)=}$19.11, p<.001) and morphological ($F_{(2,14)}$ = 23.24, p<.001) properties for Group 1. Pairwise comparisons yield a significant difference between T1 and T3 (p<.001), T2 and T3 (p<.001) for phonologically different words and a significant difference between all-time points for morphologically different words (T1 vs. T2 p=.012, T2 vs. T3 p=.001). For Group 2 there was a main effect of time for words with different phonological ($F_{(2,16)}$=32.14, p<.001) and morphological ($F_{(2, 16)}$ = 5.68 p=.014) interest. Pairwise comparisons yield a significant difference between T1 and T3 (p<.001), T2 and T3 (p=.006) for phonologically different words and a significant difference between T1 and T3 (p= 0,009) for morphologically different words.

Real word repetition

In real word repetition, analysis yield a main effect of time for Group 1 for phonological ($F_{(2,13)}$=12.10, p=.001) and morphological ($F_{(2,13)}$=7.58, p=.007) items scored for percentage of consonants correct (PCC). Pairwise comparisons yield a significant difference between T1 and T2 (p=.009), T1 and T3 (p<.001), T2 and T3 (p=.021) for phonological items, and a significant difference between T1 and T2 (p=.005), T1 and T3 (p=.004) for morphological items. For Group 2 there was a main effect of time for phonological items ($F_{(2,16)}$=7.08, p=.006) scored for PCC. Pairwise comparisons yield a significant difference between T1 and T3 (p=.004), T1 and T2 (p=.028). For morphological items, main effect of time just reached significance (p=.048) for whole word accuracy measurement, but not when PCC measurement was used.

Nonword repetition

In nonword repetition, analysis for Group 1 yield a main effect of time both for phonological ($F_{(2,13)}$=10.44, p=.002) and morphological ($F_{(2,13)}$=13.10, p<.001) items scored for PCC. Pairwise comparisons yield a significant difference on performance between all-time points for phonological (T1 vs. T2 (p=.005), T1 vs. T3 (p=.001), T2 vs. T3 (p=.045) and morphological (T1 vs. T2 (p=.001), T1 vs. T3 (p<.001), T2 vs. T3 (p=.021) items. For Group 2 there was a main effect of time for items of phonological ($F_{(2,17)}$=6.07, p=.010) and morphological interest ($F_{(2,14)}$=4.89, p=.025). Pairwise comparisons yield a significant difference on performance between T1 and T2 (p=.007), T1 and T3 (p=.026) for items of phonological interest, and between T1 and T2 (p=.019) for items of morphological interest.

Performance when stimuli of phonological and morphological interest are used

In order to explore the development of morphological elements as part of the developing speech processing system, it was explored whether linguistic domain, i.e. phonology or morphology, affects performance. Performance was compared in tasks that tap the same level of processing i.e. real word auditory discrimination and real word repetition for stimuli from these linguistic domains.

In order to explore whether the developmental pattern of real word auditory discrimination is similar both for phonological and morphological stimuli over time a 3 (Time: T1, T2, T3) by 2 (Domain: Phonological, Morphological) Repeated Measures ANOVA was performed with age group (Group 1, Group 2) as the between group factor with Bonferroni adjustment for multiple comparisons. The analyses showed a main effect of time both for Group 1 ($F_{(2,13)}$ =29.95, p<.001) and Group 2 ($F_{(2,16)}$=16.70, p<.001). There was not a main effect of linguistic domain for Group 1 ($F_{(1,14)}$=1.84, p=.197) or for Group 2 ($F_{(1,17)}$ =1.73, p=.205). The main effect of time arose because children could successfully discriminate more items over time.

In order to investigate if there is an effect of linguistic domain on real word repetition scored on a whole word basis, production of items of phonological interest and items of morphological interest were compared. A 3 (Time: T1, T2, T3) by 2 (Domain: Phonological, Morphological) Repeated Measures ANOVA was performed with age group (Group 1, Group 2) as the between groups factor with Bonferroni adjustment for multiple comparisons. The analyses showed a main effect of time both for Group 1 ($F_{(2,13)}$ =15.74, p<.001) and Group 2 ($F_{(2,16)}$=4.62, p=.026). The main effect of time arose because children could repeat more words accurately over time. There was not a main effect of linguistic domain for Group 1 ($F_{(1,14)}$=1,28, p=.419), however for Group 2 there was a just significant linguistic domain effect in favour of Phonological items ($F_{(1,17)}$=4,60, p=.047). Comparison of means did not indicate a statistically significant difference between the two tasks at any of the assessment points. As Paired Samples t-tests reveal no statistically significant difference in performance accuracy between the two tasks and the developmental pattern of the two was similar, the main effect of linguistic domain in performance of Group 2 should be treated with caution.

Performance scored for PCC was compared for items of phonological interest and items of morphological interest. A 3 (Time: T1, T2, T3) by 2 (Domain: Phonological, Morphological) Repeated Measures ANOVA, was performed with age group (Group 1, Group 2) as the between group factor with Bonferroni adjustment for multiple comparisons. The analyses showed a main effect of time both for Group 1 ($F_{(2,\ 13)}$=11.07, p<.002) and Group 2 ($F(_{2,\ 16})$=6.53, p=.008). For Group 1 there was a main effect of linguistic domain ($F(_{1,14})$=8.01, p=.013). For Group 2 group the main effect of linguistic domain missed significance ($F(_{1,17})$=3.91, p=.064).

In order to examine the relationship between performance of phonological and morphological items in real word auditory discrimination, Pearson correlations were calculated. Scores in auditory discrimination of phonological and morphological items were significantly associated within time points at T2 and T3 with a significant probability level of p< 0.033 (Bonferroni correction) for Group 1; at T1 and T3 with a significant probability level of p<0.033 (Bonferroni correction) for Group 2.

In order to examine the relationship between performance on phonological and morphological items in real word repetition, Pearson correlations were calculated within time for tasks scored for the whole word accuracy. The scoring for the overall accuracy of the word was used because in some cases the correct production of a phoneme is essential for the proper indication of a morpheme. Within time correlations between phonological and morphological items performance accuracy on real word repetition were found at all time points for both groups with a highly significant probability level of p< 0.001.

Intervention single case studies

Two children aged 4;2 and 5;7, fulfilling the same criteria for participant selection as in the normative study, who had difficulty with the production of /s/, a phoneme used in multiple phonological and morphological contexts in Greek, and fell ≤ -1.5 S.D. from age matched controls performance on several tasks, were recruited for intervention single case studies.

Four Phases of intervention (in total 24 sessions one hour sessions) were performed twice a week over the course of three months. There was an alternating focus of intervention between phonological and morphological characteristics of targets. Odd number phases targeted the production of /s/ in phonological context, in the word stem. Even number phases targeted the production of /s/ in morphological context, in word endings. Baseline assessment was carried out two months before and immediately prior to the beginning of intervention, immediately upon and two months upon completion of intervention. The test battery used in the longitudinal study was used to monitor broad changes in speech and language abilities. Upon completion of each phase of intervention, probe assessment that included some of the items treated in therapy, control and matched stimuli that carefully remained untreated, as well as more distinctive controls was conducted to monitor change and evaluate generalization.

For the youngest participant, probe assessment revealed within and across domain generalization, with gains in the production of phonological targets following morphologically oriented intervention. For the second participant, probe assessment revealed within domain gains; once /s/ was realized in a particular phonotactic structure there was generalization to other structures and control items with /z/.

Discussion

This study aimed to investigate the speech processing of phonological and morphological characteristics of words both in typically developing children and children with speech disorders. The main points arising from the normative study data indicate that tasks used are sensitive in detecting developmental progression and data can be used to evaluate performance of children with speech difficulties.

In order to explore the development of morphological elements as part of the developing speech processing system, it was investigated whether comparable speech processing skills are involved in input processing of both phonological and morphological characteristics of spoken language.

The results show that there is not a significant difference in input processing between phonological and morphological items as assessed in real word auditory discrimination.

Turning to output processing, whole word scoring did not indicate an effect of linguistic domain in Real Word Repetition for Group 1. A just significant overall linguistic domain effect in favour of phonological items was found for Group 2. However, comparison of means did not indicate a statistically significant difference between phonological and morphological items at any of the individual assessment points. Evidence for the existence of a main effect of linguistic domain in real word repetition is thus inconsistent.

Intervention case studies were carried out with two children who do not have the necessary speech processing skills for the accurate production of a phoneme that is used in morphological context. Within and across domain generalization was observed, indicating that phonological and morphological elements are part of lexical representations.

Conclusion

The present study of Greek-speaking children has contributed to our knowledge about the development of speech processing skills in the context of a highly inflectional language. Evidence was based on data from typically developing children aged 3;0-3;5 and 4;6-5;0 at the beginning of the study, assessed longitudinally on a range of tasks and from intervention single case studies of children with speech difficulties. The contribution of the present study can be summarized as follows: the evaluation of speech processing for the morphological components of language was introduced. An assessment battery was designed, which can be used as a starting point for the evaluation of speech processing skills of Greek speaking children in future studies in the field. The assessment battery may be used to diagnose children with speech difficulties in Greece.

References

Broomfield, J., & Dodd, B. (2004). Children with speech and language disability: Caseload characteristics. *International Journal of Language and Communication Disorders*, *39*(3), 303-324. http://doi.org/10.1080/13682820310001625589

Constable, A., Stackhouse, J., & Wells, B. (1997). Developmental word-finding difficulties and phonological processing: The case of the missing handcuffs. *Applied Psycholinguistics*, *18*(4), 507. http://doi.org/10.1017/S0142716400010961

Holton, D., Mackridge, P., & Philippaki-Warburton, I. (1997). *Greek: A comprehensive grammar of the modern language*. Psychology Press.

Rees, I. (2009). Deaf children's acquisition of speech skills: A psycholinguistic perspective through intervention. *Knowledge Creation Diffusion Utilization*. PhD Thesis, UCL.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, *40*(4), 708-22. http://doi.org/10.1044/jslhr.4004.708

Stackhouse, J., & Wells, B. (1997). *Children's speech and literacy difficulties: A psycholinguistic framework*. London: Whurr.

Stavrakaki, S., & Tsimpli, I.-M. (2000). Diagnostic verbal IQ test for Greek preschool and school age children: Standardization, statistical analysis, psychometric properties. In *Proceedings of the 8th Symposium of the Panhellenic Association of Logopedists* (pp. 95-106). Ellinika Grammata Athens.

Vance, M. (1997). Christopher Lumpship: Developing phonological representations in a child with an auditory processing deficit. In *Language Disorders in Children and Adults* (pp. 17–41). London, England: Whurr Publishers Ltd. http://doi.org/10.1002/9780470699157.ch2

Vance, M., Stackhouse, J., & Wells, B. (2005). Speech-production skills in children aged 3-7 years. *International Journal of Language and Communication Disorders*, *40*(1), 29-48. http://doi.org/10.1080/13682820410001716172

Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F.-X., & Lorenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proceedings of the National Academy of Sciences*, *102*(39), 14110-14115. http://doi.org/10.1073/pnas.0504446102

To cite this article: Geronikou, E., & Wells, B. (2017). Morphologically relevant speech processing skills in Greek-speaking children. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 134-140). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Mandarin lexical tone in monolingual, bilingual and trilingual children at age 2

Chang He, Ziyin Mai, Virginia Yip
changhe2015@gmail.com, maiziyin@cuhk.edu.hk, vcymatthews@cuhk.edu.hk

Childhood Bilingualism Research Centre, Chinese University of Hong Kong-Peking University-University System of Taiwan Joint Research Centre for Language and Human Complexity, University of Cambridge-Chinese University of Hong Kong Joint Laboratory for Bilingualism

Abstract. Lexical tone is one of the earliest acquired phonological features among monolingual children learning Mandarin (Hua & Dodd, 2000), yet it is one of the most challenging features for adult second-language learners (Hao, 2012; Shen, 1989). This study compares the production of Mandarin lexical tones among three children in monolingual and multilingual contexts: 1) Tong, 2;07, Mandarin monolingual; 2) Luna, 2;03, Mandarin-English bilingual; and 3) Francis, 2;05, Mandarin-English-Spanish trilingual. All three children have been exposed to Mandarin since birth. For this study, they were recorded at regular longitudinal intervals. Oral production of 44 Mandarin words was elicited through a picture-naming task modeled on that of Hua and Dodd (2000); the task involved sampling all Mandarin tones and phonemes in each legal word position. We compared the children's performance on the four basic lexical tones across 56 syllables and corroborated the results with data from the longitudinal recordings. The data show that the mid-rising and falling-rising tones were more challenging for the bilingual and trilingual children than for their monolingual peer. Delay and transfer occur in the bilingual and trilingual children's tone acquisition. The trilingual child may have applied bilingual optimization strategies. In addition, this study demonstrates that the children acquired high-level and high-falling tones before mid-rising and falling-rising tones (Li & Thompson, 1977). Overall, Mandarin lexical tone in multilingual children can be characterized by the early emergence of tonal categories as well as by variability and vulnerability in fine acoustic features (which is attributable to cross-linguistic influence from nontonal languages).

Keywords: Mandarin tone, language acquisition, bilingual and trilingual children, child phonology, multilingualism

Introduction

Lexical tone is one of the earliest acquired phonological features among monolingual children learning Mandarin (Hua & Dodd, 2000), yet it is one of the most challenging features for adult second-language learners (Hao, 2012; Shen, 1989). In Hua and Dodd's (2000) study, Mandarin-speaking children acquired tone first, followed by syllable-final consonants and vowels, and finally syllable-initial consonants. Monolingual children acquire the tonal system by age two. However, research on multilingual children's acquisition of Mandarin lexical tone is very rare. This study compares the production of Mandarin lexical tones among three children in monolingual and multilingual contexts: Tong (Mandarin monolingual), Luna (Mandarin-English bilingual) and Francis

(Mandarin-English-Spanish trilingual). All three children have been exposed to Mandarin since birth. For this study, we recorded them at regular longitudinal intervals. We also elicited oral production of 44 Mandarin words through a picture-naming task modeled on that of Hua and Dodd (2000); the task involved sampling all Mandarin tones and phonemes in each legal word position. We compared the children's performance on the four basic lexical tones across 56 syllables and corroborated these results with data from the longitudinal recordings. Our data show that early establishment of tonal categories extends to children learning Mandarin in bilingual and trilingual contexts. However, our acoustic analysis reveals between-child differences in terms of pitch range, with an inverse relationship between number of languages and pitch range.

Acquisition of Mandarin lexical tone

Recent studies have concentrated on the phonological acquisition of adult second-language speakers (e.g., Best & Tyler, 2007; Flege, 2007). Little attention has been paid to the phonological acquisition of simultaneous bilingual and trilingual children. However, many simultaneous bilingual and trilingual children are acquiring Mandarin throughout the world. By exploring Mandarin tone production, this paper sheds light on Mandarin lexical tone acquisition in monolingual and simultaneous bilingual children.

The phonologies of a multilingual speaker's languages are separate but interdependent (Fabiano-Smith & Goldstein, 2010; Kehoe, 2002). Interaction appears in three ways: acceleration, delay and transfer (Paradis & Genesee, 1996). Acceleration and delay refer to multilinguals' rate of acquisition in comparison to monolinguals', and transfer refers to the use of a specific linguistic property of one language in another language (Keffala, Barlow, & Rose, 2016). In one study, Mandarin-English bilingual toddlers showed reduced sensitivity to tone mispronunciations compared to their Mandarin monolingual peers; this was attributed to the influence of English (Wewalaarachchi, Wong, & Singh, 2017).

Mandarin has four basic lexical tones: high-level (Tone 1), mid-rising (Tone 2), falling-rising (Tone 3) and high-falling (Tone 4) (Duanmu, 2004). The present study adopted Chao's (1930) method to mark Mandarin lexical tones, in which tone letters range from 1 (low pitch, or F0, within the speaker's range) to 5 (high F0). Additionally, this study used Praat to confirm the objectivity of the tones. This study addresses the following research questions:

1. How does the acquisition of Mandarin lexical tone differ among monolingual, bilingual and trilingual children?

2. What are the characteristics of multilingual children's tone systems?

3. What forms of interaction take place in multilingual development?

Methodology

We analyzed three children who had been exposed to Mandarin since birth in monolingual (Tong, Mandarin), bilingual (Luna, Mandarin-English) and trilingual (Francis, Mandarin-Spanish-English) contexts. Table 1 shows their linguistic backgrounds.

Table 1. Backgrounds of target children

	Tong (Monolingual)	Luna (Bilingual)	Francis (Trilingual)
Gender	M	F	M
Language exposure (fraction and source of language input)	Mandarin (All)	Mandarin (Parents, 1/2) English (School, 1/2)	Mandarin (Mother, 1/3) Spanish (Father, 1/3) English (School, 1/3)
Residence	Shenzhen, China	Maryland, USA	Hong Kong

Tong's tone production is used as a baseline in this study. Luna, a heritage child and Mandarin-English bilingual who was born and raised in the US (second-generation immigrant), received Mandarin input at home and English input at nursery school. At home, Francis was addressed in Mandarin by his mother and in Spanish by his father; at nursery school, he was exposed to English.

We conducted a picture-naming task modeled on that of Hua and Dodd (2000) with these target children to collect spontaneous data on their tone production. The words tested included 44 words (39 nouns, 4 verb phrases, and 1 color term) that young Mandarin-speaking children are likely to know. The experiment was videotaped and occurred in the presence and with the guidance of a parent. Most productions were spontaneous naming with few that the children could not name, repetitions were solicited. Each recording was 10 to 15 minutes long. We gave each child at least three chances to pronounce each word to confirm the child's pronunciation.

To examine whether the data from the longitudinal recordings confirmed the experimental data, we also extracted some tone-production data from the corpus for acoustic analysis (see Table 2).

Table 2. Corpus details

	Tong	Luna	Francis
Age range	1;07-3;04	2;00-4;11	2;00-2;11
Language(s)	Mandarin	Mandarin and English	Mandarin, English and Spanish
Number of recordings	22	20	11 (as of August)
Hours of recordings	22	25	11 (as of August)

For each child, we selected a recording that we had made around one month before the experiment and another that we had made around one month after the experiment. We then selected five words for each tone to compare the children's tone production across the experimental and corpus data. Thus, there were 40 words in total for each child, 10 for each tone. Most selected words were located in the initial position of two-word or three-word utterances so as to avoid phonetic influence from other words. As with the experimental analysis, we did not include words affected by sandhi and/or light tones.

The present study adopted Chao's (1930) five-point notation to mark Mandarin lexical tones, in which the four tones are represented as [55] (Tone 1, high-level), [35] (Tone 2, mid-rising), [214] (Tone 3, falling-rising), and [51] (Tone 4, high-falling). As the light tone is subject to regional differences, we did not include it in this study. We eliminated 16 syllables affected by light tones (e.g., *yang* in *tàiyang*, meaning "sun") or sandhi tones (e.g., *shóu* in *shóuzhǐ*, meaning "finger"). We then compared the three children's performance on the four basic lexical tones across 56 syllables. We used Praat to extract the parts needed for analysis and ProsodyPro (a Praat script) to generate the values for each syllable's fundamental frequency, F_0 (Xu, 2013). After calculating the average F_0 values of each point for all syllables using the same tone, we applied the z-score (LZ-score) transformation (Zhu, 2004) to obtain normalized F_0 data. With these logarithmic F0 values, we drew 10-point pitch graphs for the four tones.

In addition to analysis using Praat, we also asked two adult native Mandarin speakers to judge the tones the children produced. The adults were given the selected 56 syllables without information about the original words or tones. They then determined which tones they heard in each case.

Results

Figure 1 presents the target children's experimental pitch patterns. Tong's tone contours show the typical characteristics of Mandarin tones (Chao, 1930; Li & Thompson, 1977). Luna's tone system maintains the four-way distinction, but the contours are generally flatter than those in Tong's system. Compared to Tong's tones, Luna's tone 2 begins lower, and her tone 4 ends higher. Francis' Mandarin tone system is qualitatively different, indicating a tendency to merge both tones 1 and 4 and tones 2 and 3.

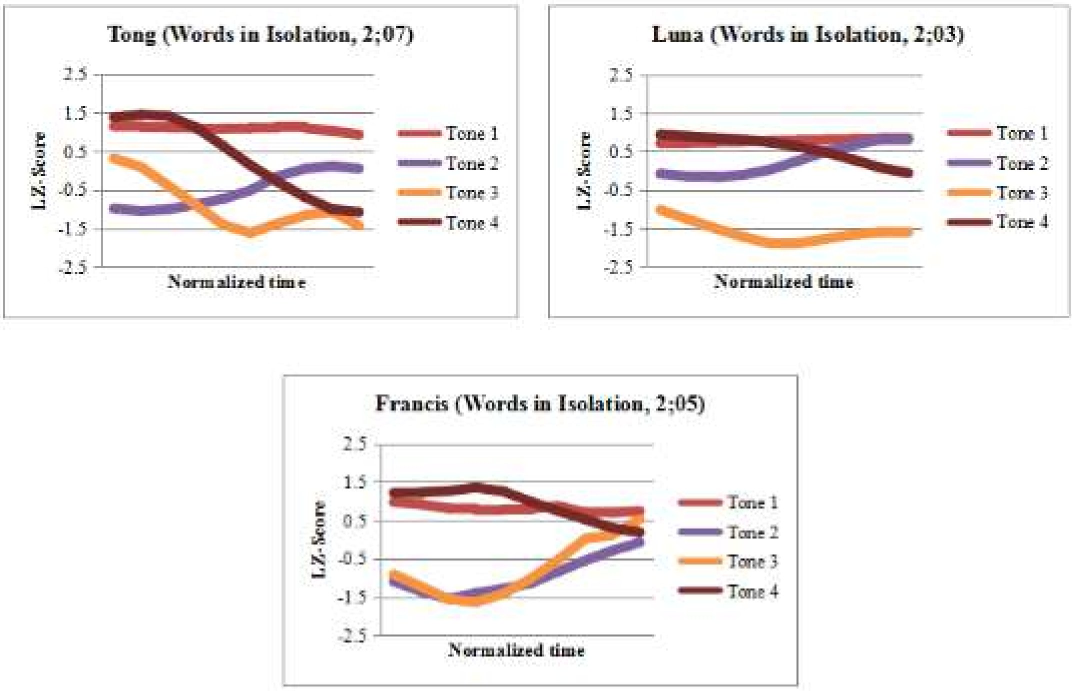


Figure 1. Tone frequencies of words in isolation

Figure 2 shows the pitch patterns of the target children in the longitudinal recordings. Tong maintains 4 tonal categories in running speech. However, Luna's tone 2 and tone 4 deviate from the baseline. Except for tone 4, Francis' tones are not target-like, and tone merging is not seen in his corpus data. Generally speaking, the corpus data of the three subjects confirms the experimental findings.

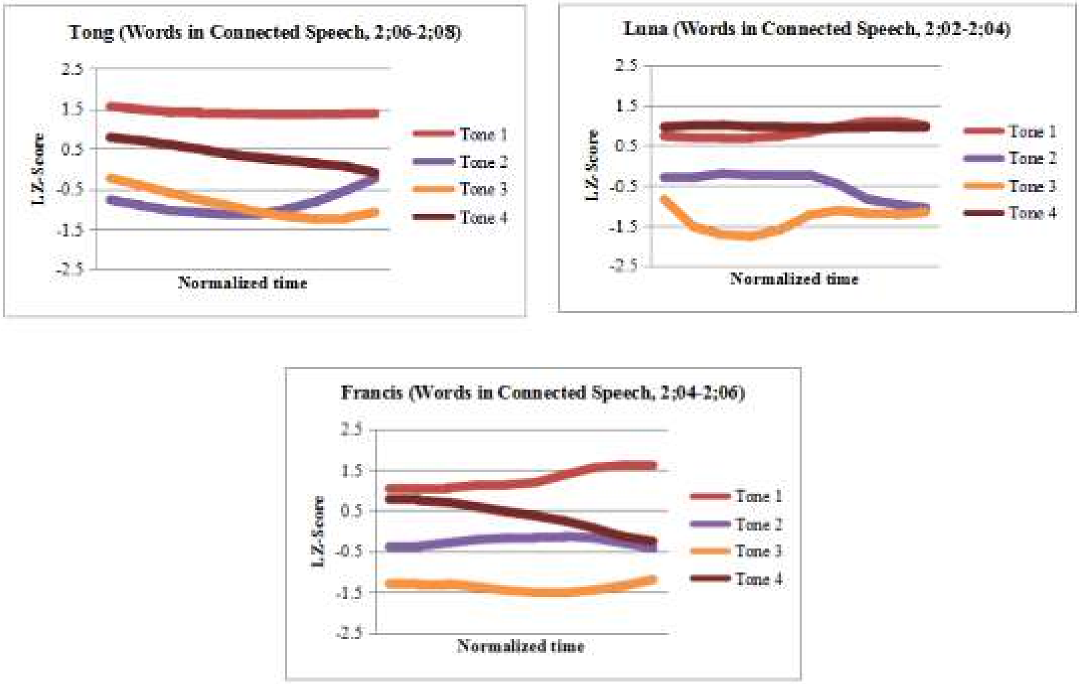


Figure 2. Tone frequencies of words in connected speech

Based on the native speakers' judgments, Luna's performance is comparable to Tong's (see Table 3). However, Francis' tones are far from the target tones, especially for tone 3, none of the words for which were perceived as being tone 3 in the adult native speakers' judgments. Most of Tong and Luna's four tones are accurately recognized by both native speakers.

Table 3. Native Speaker Judgment

Tone	Total No. of tokens	Number and percentage of tokens that native speakers perceived as accurate					
		Tong		Luna		Francis	
1	Experiment: 18	13	72%	12	67%	7	39%
	Corpus: 10	5	50%	4	40%	7	70%
2	Experiment: 16	7	44%	8	50%	7	44%
	Corpus: 10	3	30%	0	0%	2	20%
3	Experiment: 12	5	45%	6	50%	0	0%
	Corpus: 10	3	30%	2	20%	0	0%
4	Experiment: 10	6	60%	6	60%	3	30%
	Corpus: 10	3	30%	2	20%	3	30%

*: Tokens with at least 40% recognition are highlighted in gray

Discussion

The three children's tone performance can be ranked in descending order of number of languages spoken, from monolingual to trilingual. The monolingual child established all four tonal categories in both isolated words and connected speech. The bilingual child's four tones in the experimental data were comparable to those of the monolingual child, except with flatter contours. The mid-rising tone seemed to be most challenging to the bilingual child. The trilingual child's tonal system was least target-like, with an emerging 2-tone system.

Our data support Li and Thompson's (1977) observation that the high-level and high-falling tones are acquired before the mid-rising and falling-rising tones. The bilingual child showed difficulty acquiring the rising tone, and the trilingual child had difficulty acquiring the rising and falling-rising tones. Moreover, the trilingual child's mid-rising and falling-rising tones were acoustically similar, suggesting a tone merger in that child's phonological system.

Influence from nontonal languages may account for the non-target-like tone contours observed bilingual and trilingual children. In this case, English and Spanish affected Mandarin. The tone production of the bilingual and trilingual children resembled the patterns shown in adult second-language learners of Mandarin whose native languages were nontonal (Broselow, Hurtig, & Ringen, 1987).

In the trilingual child's tone acquisition, the child may have applied simplification and bilingual optimization strategies (Muysken, 2013). He might have matched patterns from Mandarin to those of the other two languages when possible. If so, this adaptation may have been driven by the child's knowledge of two nontonal languages (English and Spanish) that rely on the rising-falling contrast to differentiate pragmatic meanings (such as interrogative and non-interrogative).

Conclusion

This study demonstrates that the mid-rising and falling-rising tones are challenging to bilingual and trilingual children who are acquiring Mandarin. Transfer may take place in their tone acquisition, which differs from that of their monolingual peers, who typically acquire Mandarin tone features by age two. The trilingual child might have applied bilingual optimization strategies. This study also provides evidence that Mandarin's mid-rising and falling-rising tones are acquired after its high-level and high-falling tones. Overall, Mandarin lexical tones in multilingual children are characterized by the early emergence of tonal categories and by the variability and vulnerability of the fine acoustic features (which is attributable to cross-linguistic influence from nontonal languages).

References

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M.J. Munro & O. S. Bohn (eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Broselow, E., Hurtig, C. R., & Ringen, C. (1987). The perception of second language prosody. In G. Ioup and S. H. Weinberger (eds.), *Interlanguage Phonology* (pp. 350-362). Cambridge, MA: Newbury House.

Chao, Y. R. (1930). A system of tone letters. *La Maître Phonétique*, *45*, 24-27.

Duanmu, S. (2004). Tone and non-tone languages: An alternative to language typology and parameters. *Language and Linguistics*, *5*(4), 891-923.

Fabiano-Smith, L., & Goldstein J. A. (2010). Phonological Acquisition in Bilingual Spanish–English Speaking Children. *Journal of Speech, Language, and Hearing Research*, *53*(1), 160-178. doi:10.1044/1092-4388(2009/07-0064)

Flege, J. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & Hualde, J. (eds.), *Laboratory Phonology 9* (pp. 353-380). Berlin: Mouton de Gruyter.

Hao, Y. C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, *40*, 269-279.

Hua, Z., & Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language*, *27*, 3-42.

Keffala, B., Barlow, J. A., & Rose, S. (2016). Interaction in Spanish-English bilinguals' acquisition of syllable structure. *International Journal of Bilingualism*. July 15, 2017 Online: https://doi.org/10.1177/136700 6916644687.

Kehoe, M. (2002). Developing vowel systems as a window to bilingual phonology. *International Journal of Bilingualism*, *6*(3), 315-334.

Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, *4*(2), 185-199.

Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and cognition*, 16, 709-730.

Paradis, J., & Genesee, F. (1996). Syntactic acquisition in bilingual children: Autonomous or interdependent? *Studies in Second Language Acquisition*, *18*(1), 1-25.

Shen, X. N. (1989). Americans learning Mandarin tones. *Chinese Teaching in the World*, *3*, 158-168.

Wewalaarachchi, T. D., Wong, L. H., & Singh, L. (2017). Vowels, consonants, and lexical tones: Sensitivity to phonological variation in monolingual Mandarin and bilingual English-Mandarin toddlers. *Journal of Experimental Child Psychology*, *159*, 16-33.

Xu, Y. (2013). ProsodyPro - A tool for large-scale systematic prosody analysis. *Tools and Resources for the Analysis of Speech Prosody*, 7-10.

Zhu, X. N. (2004). Normalization of F_0. *Language Science*, *2*, 3-19.

To cite this article: He, C., Mai, Z., & Yip, V. (2017). Mandarin lexical tone in monolingual, bilingual and trilingual children at age 2. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 141-147). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Misunderstanding as a two-way street: Communication breakdowns in native/non-native English/French tandem interactions

Céline Horgues, Sylwia Scheuer
celine.horgues@univ-paris3.fr, sylwia_scheuer@yahoo.fr

Université Sorbonne Nouvelle-Paris 3

Abstract. The paper explores the bilingual data contained in the SITAF tandem corpus, collected at the Sorbonne Nouvelle University. As part of the SITAF project (Horgues & Scheuer, 2015), we gathered video and audio recordings from face-to-face conversational exchanges held by 21 pairs of undergraduate students, with each such 'tandem' consisting of a native speaker (NS) of English and a native speaker of French. The tandems performed collaborative semi-spontaneous conversation tasks (story-telling and debating) in both languages. With its largely unscripted L1-L2 productions, the corpus offers ample opportunities for various types of analyses of NS-NNS interactions, including studies of communication breakdowns. We will be looking at both sides of the communicative picture, i.e. cases where it was the nature of either the native, or the non-native participant's output that seemed to be the main stumbling block, with a view to reporting on the number and the linguistic sources of misunderstandings attested in both English and French conversations. The paper will show how these issues are resolved in a highly empathetic and collaborative manner, in keeping with the tandem learning environment, where the relation between the participants is relatively non-hierarchical and reversible.

Keywords: tandem learning, native and non-native speakers, English/French, communication breakdowns, intercultural interaction

Introduction

The traditional take on interactions between L1 and L2 speakers, which prevailed at least until the late 20th century, construed intelligibility as 'a one-way process in which non-native speakers are striving to make themselves understood by native speakers whose prerogative it was to decide what is intelligible and what is not' (Bamgbose, 1998:10). This view is nowadays generally regarded as indefensible, as it ignores the L1 speaker's potential for not being understood, and also fails to assign a fair portion of responsibility for communication breakdowns to the receiver. In accordance with the Gricean Cooperative Principle, successful communication 'rests on mutual *trust* between *responsible* individuals', with the addressees adopting therefore a 'benevolent' or 'charitable' attitude" (Dascal, 1999:757), which allows them to ignore certain imperfections in the speakers' output in a bid not to disrupt the flow of the conversation.

We follow Bamgbose's (1998:11) definition of intelligibility as 'a complex of factors comprising recognizing an expression, knowing its meaning, and knowing what that meaning signifies in the sociocultural context'. Misunderstanding has been variously defined, sometimes quite broadly e.g., by Mauranen (2006:128) as 'a potential breakdown point in conversation, or at least a kind of communicative turbulence', while some linguists distinguish it from miscommunication, viewing the latter as unresolved, or sustained, misunderstanding (Dascal, 1999). In the present paper, we adopt the all-embracing approach to misunderstanding, using the term interchangeably with 'miscommunication' or 'communication breakdown', and taking it to also include non-understanding (see Jenkins, 2000). We thus employ it to denote all cases where the listener demonstrably has difficulty, or is incapable of, grasping the meaning of an utterance as, seemingly, intended by the speaker, notwithstanding the fact that what the speaker truly means is often a matter of speculation.

Even though human communication is an endeavour where mutual intelligibility between interlocutors is the default assumption (Mauranen, 2006:123), it does nevertheless rely on inferences and guesswork that may misfire, meaning that '[t]he possibility of misunderstanding is ever present because not everything can be explicitly said' (Dascal, 1999:755). Communicative turbulence is almost certainly more widespread than is overtly manifested, considering that both parties approach the interaction in a relatively egocentric manner (Keysar, 2007), with speakers routinely believing that their message is accurately understood by the addressee more often than it really is. Given that misunderstandings are rooted in limited common background, or shared 'belief space', between the interlocutors, 'the less interlocutors know about each other, the more likely they are to misunderstand each other on a linguistic, social, or cultural level' (Varonis & Gass, 1985:327). Consequently, it comes as no surprise that there are more indications of communication breakdowns in NS-NNS (native speaker – non-native speaker) than in NS-NS conversations, as native and non-native speakers 'may have radically different customs, modes of interacting, notions of appropriateness, and, of course, linguistic systems', which renders them 'multiply handicapped' in interactions with one another (Varonis & Gass, 1985:327, 340). In terms of the major difficulties inherent to NNS output, one can naturally invoke their insufficient mastery of the linguistic system (however defined), which may result in what Varonis and Gass (1985:334) term 'noise' in the speaker's utterance, produced by e.g., accent or ungrammaticality. NS discourse, on the other hand, may present processing challenges of its own by virtue of showing insufficient accommodation to the needs of the NNS interlocutor. Embracing these needs ideally means avoiding 'slang, opaque idioms, rapid speaking rates, and culture-specific references' (Trudgill, 2005: 82). Just like misunderstanding is a multifaceted product of speaker-listener interaction, so is its management. Resolving misunderstandings is a collaborative process involving negotiating and co-constructing meaning, and culminating in the interactants 'coming to an understanding' (Weigand, 1999:766), even if the strategies for achieving this goal tend to vary depending on whether or not the participants share their mother tongue (Wong, 2000). This expected cooperation and reaching out are natural corollaries of the communicators being bound by 'the duty to make oneself understood and the duty to understand' (Dascal, 1999:757) and, consequently, the 'duty' to embark on a joint trouble-shooting mission if need arises.

The tandem setting or, the tandem 'constellation' (Kohn, 2016), is an environment where principles of charity, collaboration, reciprocity and solidarity between the communicators certainly come to the fore (Brammerts & Calvert, 2003). O'Rourke (2005:434) defines tandem learning as 'an arrangement in which two native speakers of different languages communicate regularly with one another, each with the purpose of learning the other's language', with each participant assuming both the role of the learner and the role of the expert (although at different points in the interaction), thus making their relationship essentially non-hierarchical. It is important to emphasize that this context is therefore different, despite areas of overlap, not only from the lingua franca one (due to the learner's conscious orientation towards the native speaker model), but also from the 'classic' NS-NNS setting where roles are not reversible. As a result, there is reason to expect misunderstandings not only to be resolved in a particularly cooperative fashion, but also to be less likely to arise in the first place. Since the experience of being both the NS and the NNS part of the dialogue is never far off, tandem partners may be very proactive in averting communication breakdowns, for example by maximizing simplicity of expression or making ostensibly excessive use of confirmation checks, clarifications and repetitions (cf. Mauranen, 2006).

Our objective is to provide corpus-based data on the frequency of overtly marked communication breakdowns in English and French, the relative contribution of the NS vs the NNS output to generating misunderstandings, the types of linguistic sources of the problems, and the various multimodal ways in which the tandem partners signal and resolve communication issues.

Method

The SITAF corpus (Horgues & Scheuer, 2015), which serves as our database, consists of around 25 hours of audio- and video-recorded, face-to-face interactions held in English and French by 21 pairs of native-French speaking (coded F01-F21) and native-English speaking (coded A01-A21) tandem participants, all of whom were students at the Sorbonne Nouvelle University, aged 17-22. Prior to participating in the tandem programme, the candidates filled in an online questionnaire where, among other things, they were asked to self-assess their mean proficiency and oral expression in the L2 (English or French, as appropriate). The averaged scores were, respectively, 7.2/10 and 6.8/10 for the Francophones, and 6.9/10 and 6.6/10 for the Anglophones. It is only the between-group difference in the latter score (oral expression) that reached statistical significance.

The speakers were recorded on two occasions: sessions 1 and 2, separated by a 3-month interval, while performing three types of tasks. Two of them were communication activities, *Liar-Liar* (game 1; storytelling) and *Like Minds* (game 2; debating), while the last was a monitored reading task. Although all the participants performed all three tasks in their respective L1 and L2 at least once during the recording sessions, including control L1-L1 interaction, our present study will only be concerned with 'session 1, game 2' portion of the data. Game 1 was deemed less suitable for this analysis by virtue of according disproportionately more speaking time to one of the participants (the storyteller), which might have skewed the overall results.

Reference will also be made to the metadata collected through the questionnaires filled in by the participants on completing the programme (one regarding their overall tandem experience and another specifically focusing on pronunciation issues), with a view to comparing their introspective judgements, as 'misunderstanders' and 'misunderstandees', with the corpus-derived data.

Results

We analysed 42 video sequences from the SITAF corpus (session 1, game 2): 21 interactions in French and 21 interactions in English performed by the same tandem pairs. Out of the 2 hours and 30 minutes of video-recorded speech studied, we identified 39 cases of detectable misunderstanding, which were distributed as follows: 13 cases in French (no case detected in 11 conversations) and 26 cases in English (no case detected in 8 conversations only). The difference between the two language conditions is statistically significant ($p<.04$).

Participants' roles in misunderstanding management

In the 39 cases of misunderstandings, the NNS participant was the main speaker (whose speech was misunderstood) in about two thirds of cases (16/26 in the English conversations and 8/13 in the French conversations, see Table 2). Thus in the remaining cases, the NS participant's output was misunderstood, proving the point made in the title of this paper: misunderstanding is not a one-way street restricted to non-native speech.

In both sets of conversations, misunderstandings were most frequently detected and signaled by the recipient (two thirds of cases). Occasionally, both participants simultaneously spotted the problem. In line with what has been presented in the literature (Dascal, 1999), most misunderstandings were detected instantaneously. Cases of delayed detection and resolution were relatively rare (6/39) but are worth further study. Cases of misunderstandings going unnoticed were very rare in our data (2/39).

Frequency of misunderstandings

The imbalance observed in the number of misunderstandings between the French and English conversations (twice as many in the latter) was completely unexpected, especially since the research into corrective feedback instances we previously conducted (Scheuer & Horgues, 2016) showed the opposite trend. Indeed, we found that twice as much corrective feedback was provided by Francophone speakers in the French conversations compared to what their Anglophone counterparts did in the English conversations. We will revisit this discrepancy in the discussion section.

Since we wanted to compare the actual observation of misunderstandings with the participants' introspective representation of this same phenomenon, we then inspected the post-recording metadata questionnaires. Table 1 presents a summary of these results.

The tandem participants were asked to estimate the frequency at which their partner's L2 pronunciation impeded their comprehension (the middle column). On average, the Francophone participants estimated that their partner's L2 French pronunciation caused comprehension problems

11.9% of the time and the Anglophones estimated it was the case for their partner's L2 English pronunciation 7.9% of the time (the difference is statistically significant; $p<.05$). Another question investigated the self-assessed frequency at which they *thought* their *own* L2 pronunciation caused comprehension problems for their interlocutor (the last column). They estimated that their pronunciation in L2 made them unintelligible 15.6% of the time (Francophones in English) and 18.1% of the time (Anglophones in French).

Table 1. Self-estimated frequency of misunderstandings due to L2 pronunciation

Self-assessment by	*Misunderstands partner*	*Gets misunderstood by partner*
Francophones	11.9% time	15.6% time
Anglophones	7.9% time	18.1% time

Factors of misunderstandings

Most triggers causing the detectable misunderstandings consisted of purely linguistic factors. The communicative environment and nature of the task certainly reduced the potential for miscommunication pertaining to the non-literal meaning (rhetorical/pragmatic and intercultural triggers). Still, some comprehension difficulties were caused by different socio-cultural assumptions (e.g., the cost of tuition fees in France vs. the USA).

Table 2. Types of triggers in the English and French conversations

	ENG. conversations		**FR. conversations**	
Triggers	NS speech	NNS speech	NS speech	NNS speech
Lexical	*4*	*5*	*5*	*3*
Phonetic	*2*	*7*	*0*	*4*
Grammar/morphosyntax	*0*	*1*	*0*	*0*
Pragmatic/rhetorical	*3*	*1*	*0*	*0*
Combined	*1*	*2*	*0*	*1*
TOTAL	***10***	***16***	***5***	***8***
	26		***13***	

The figures in Table 2 show that:

- Most frequently, misunderstandings were due to lexical triggers (incorrect/unknown lexeme, collocations or idiom) e.g., the English compound *tuition fees* or the French colloquial expression *une boîte à fric* 'a scam'.
- Phonetic triggers were just as frequent as lexical triggers in the English conversations, and slightly less frequent in the French conversations.

- Morphosyntactic triggers were extremely rare in the English conversations (e.g., incorrect sentence structure in: F06 ‘well, it depends on erm how many the prisoners would stay in jail...’ A06 ‘How…sorry?’ [later] ‘Oh! how long would the prisoners be in jail for?’.

More specifically, the phonetic triggers observed in the French conversations related to the pronunciation of final vowels or the confusion between the typical high front-rounded French vowel /y/ and other similar vowels like /i/, or more globally to choppy or disfluent phrasing. In the English conversations, the most frequent triggers related to word stress placement (critical for French learners since word stress is not distinctive in their mother tongue) in words like *tu'ition, 'citizen* or *'prisoners,* but also to some vowels and consonantal phonemes absent from the French inventory (e.g., the interdental fricative <th> as in *thief*).

Again, the post-recording questionnaires were useful in providing metadata about what the participants *attributed* the misunderstandings to. In their opinion, lexical issues (more than phonetics) were the main cause. Interestingly, they made a connection between moments when they lacked or were uncertain about vocabulary and the reinforcement of their foreign accent. More marginally, they quoted sentence structure and psychological issues like anxiety and shyness as possible causes.

Strategies for dealing with misunderstandings

More often than not, the participants were proactive in using a combination of non-mutually exclusive strategies for managing and solving misunderstandings, such as:

- Verbal strategies with comprehension checks by the interlocutor (asking clarification questions e.g., *the what*?*)* or by the main speaker anticipating comprehension issues (e.g., *do you understand the topic ok*?), or with reformulations (switching language as a last resort).
- Vocal strategies to signal or resolve comprehension issues: increasing the frequency and length of pauses, using rising intonation, decreasing tempo, hyperarticulating.
- Non-verbal strategies, especially through face expressions (frowns, squints), gaze direction shifts, head movements for backchannelling, hand movements to back up clarification and reformulation. Interestingly, trunk orientation also supported attempts by the interlocutor to ‘reach out’ or ‘come’ to an understanding with their partner.

Undoubtedly, the easiest, face-saving strategy was simple avoidance, which is impossible to objectively quantify but is illustrated by a quote from an Anglophone participant (A15): ‘When I don't understand/there are problems of understanding, I normally try to move past them and ignore them to continue the flow of the conversation’.

Discussion

We had not predicted a difference to arise in the number of misunderstandings depending on the language being spoken. A range of hypotheses are worth exploring to account for this observation. It is important to emphasize that the difference in the (self-assessed) L2 proficiency level between the two sets of participants was not sufficient to account for such a discrepancy. First, the French conversations turned out to be slightly shorter (mean length: 3 mins 11 s) than the English conversations (4 mins). This difference in mean length was very nearly large enough to reach statistical significance ($p<.057$), but the extra minute of speaking time in English cannot solely account for misunderstandings being twice as frequent. When interacting in French, the Anglophone participants might have been more polite about incomplete clarity of expression in their partner’s L1 speech (see the sociocultural differences underlined in Scheuer & Horgues, 2016), and therefore more reluctant to signal misunderstanding than their French counterparts in the English part of the data. Finally, due to the Anglophones’ ongoing immersion in L2 communication on a daily basis (their residency in France at the time of the experiment), mutual understanding might simply have been better in French and the Anglophones might have developed more expert strategies for anticipating and avoiding misunderstandings than their Francophone partners.

What emerges from the tandem partners' self-estimations of the relation between L2 pronunciation and (mis)understanding is that both sets of participants believed that their *own* L2 pronunciation caused more misunderstandings than *their partner's* L2 pronunciation (Table 1). This trend reaches statistical significance for the Anglophones ($p<.006$) but not for the Francophones. Overall, they all tended to be harsher on themselves than on their partners when taking the blame for causing misunderstandings in tandem interactions. Both sets of participants also largely underestimated how intelligible they were in L2 compared to what their interlocutor declared (statistically significant; $p<.05$ for ANGs and $p<.0004$ for FRs).

The confrontation of the metadata questionnaires and actual observation of sequences shows the participants' awareness about miscommunication factors and the link between pronunciation and intelligibility. Indeed, their evaluation of L2 pronunciation features impairing comprehension was in keeping with our analysis of misunderstanding sequences. Phonetic triggers were the most frequent in the NNS-generated output, which lends support to Pickering's (2006:222) observation that 'undoubtedly the most salient speaker factor in standard NS-NNS comprehensibility studies has been phonology and primarily the effects of accentedness'.

Conclusion

Our subjects were found to be quite proactive about anticipating, preventing and resolving misunderstandings, but at the same time they also proved relatively tolerant and empathetic when they sometimes preferred to ignore comprehension issues not to disrupt the flow of the conversation and not to cause embarrassment (face-loss) to themselves or the partner. Most cases of misunderstandings were triggered by lexical and phonetic issues in the output of NNS participants but also in the speech of NS participants, revealing a mismatch with (and sometimes lack of accommodation to) their interlocutor's competence.

Our face-to-face interactions emphasized the role of a whole range of multimodal strategies to manage misunderstandings in tandem learning. It is particularly enlightening to notice how misunderstanding management is a collaborative and negotiation process to which both participants contribute.

We might wonder whether the solidarity and empathy at the basis of tandem principles might have made our participants over-cooperative or, to take Varonis and Gass (1985)'s terms, excessively 'charitable' in the way they managed miscommunication.

Among future perspectives we hope to compare the present results with misunderstanding management in NS-NS (control) interactions and we would also like to examine how growing familiarity between tandem participants affects the frequency of detectable misunderstandings by taking the longitudinal perspective (same participants after 3 months of tandem practice – game 2, session 2).

References

Bamgbose, A. (1998). Torn between the norms: innovations in world Englishes. *World Englishes*, *17*(1), 1-14.

Brammerts, H., & Calvert, M. (2003). Learning by communicating in tandem. In Lewis, T. & L. Walker (eds.), *Autonomous language learning in tandem*. Sheffield, UK: Academy Electronic Publications.

Dascal, M. (1999). Introduction: Some questions about misunderstanding. *Journal of Pragmatics*, *31*, 753-762.

Horgues, C., & Scheuer, S. (2015). Why some things are better done in tandem? In J. A. Mompeán & J. Fouz-González (eds.), *Investigating English Pronunciation: Current Trends and Directions* (pp. 47-82). Basingstoke and New York: Palgrave Macmillan.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford, UK: Oxford University Press.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics*, *4*(1), 71-84. Retrieved 28 Aug. 2017, from doi:10.1515/IP.2007.004.

Kohn, K. (2016). From ELF communication to lingua franca pedagogy. In Pitzl, M.-L. & R. Osimk-Teasdale (eds.), *English as a lingua franca: Perspectives and prospects. Contributions in Honour of Barbara Seidlhofer* (pp. 87-96). Boston/Berlin: De Gruyter Mouton.

Mauranen, A. (2006). Signaling and preventing misunderstanding in English as a lingua franca communication. *International Journal of the Sociology of Language*, *177*, 123-150.

O'Rourke, B. (2005). Form-focused interaction in online tandem learning. *CALICO Journal*, *22*(3), 433-466.

Pickering, L. (2006). Current research on intelligibility in English as a lingua franca. *Annual Review of Applied Linguistics*, *26*, 219-233.

Scheuer, S., & Horgues, C. (2016). On the (a)symmetry of FR-ANG-lais pronunciation feedback: a comparative analysis of corrective feedback provided by English-speaking and French-speaking participants during language tandem conversations. Book of Abstracts of the *8th International Symposium on the Acquisition of Second Language Speech*, *New Sounds 2016*. University of Aarhus, 164.

Trudgill, P. (2005). Native-speaker segmental phonological models and the English Lingua Franca Core. In K. Dziubalska-Kołaczyk & J. Przedlacka (eds.), *English Pronunciation Models: A Changing Scene. A Democratic Accent for the Global Village?* (pp. 77-98). Frankfurt a/M: Peter Lang.

Varonis, E. M., & Gass, S. (1985). Miscommunication in native/nonnative conversation. *Language in Society*, *14*(3), 327-343.

Weigand, E. (1999). Misunderstanding: The standard case. *Journal of Pragmatics*, *31*, 763-785.

Wong, J. (2000). Delayed next turn repair initiation in native/non-native speaker English conversation. *Applied Linguistics*, *21*(1), 244-267.

To cite this article: Horgues, C., & Scheuer, S. (2017). Misunderstanding as a two-way street: Communication breakdowns in native/non-native English/French tandem interactions. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 148-154). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Phonology of late and precocious talkers in French

Margaret Kehoe, Tamara Patrucco-Nanchen
margaret.winkler-kehoe@unige.ch, tamara.patrucco@unige.ch

University of Geneva

Abstract. This study examines the relationship between lexical and phonological variables in 40 French-speaking children, aged 2;5. Specifically, it examines the influence of phonetic complexity (PhC), phonological production (PhP), and neighbourhood density (ND) on vocabulary size. Children were divided into four groups on the basis of vocabulary size: late1 (<10%ile), late2 (15-25%ile), middle (40-60%ile), and precocious (>90%ile). The children's lexicons were coded in terms of PhC and ND (one-syllable words), and their PhP capacities were determined from measuring percent consonants correct (PCC) in spontaneous language samples. Results indicated significant group differences in all three variables. Children with larger vocabularies selected words with greater PhC and with lower ND values. They had superior PhP abilities compared to children with smaller vocabularies. Linear regression indicated that 71% of variance in vocabulary size could be accounted for by the three variables, with the highest percentage accounted for by ND (56%). Our findings are consistent with previous studies which show that ND plays an important role in accounting for variance in vocabulary size. They also indicate that other variables such as PhC and PhP influence lexical acquisition.

Keywords: lexical selection, phonetic complexity, neighborhood density, phonological production, late talkers

Introduction

The relationship between lexical and phonological development has been the subject of much research in recent times (see Stoel-Gammon, 2011, for a review). One line of research has focused on children's tendency to select and avoid words on the basis of their phonological characteristics, a phenomenon referred to as *lexical selection and avoidance*. Another line of research has examined the relationship between phonological production (PhP) and vocabulary size (Rescorla & Ratner, 1996; Smith, McGregor, & Demille, 2006). Yet another line of research has adopted variables from adult psycholinguistics such as neighbourhood density (ND) and word frequency (WF) to determine what factors account for vocabulary development in children (Stokes, 2010; Stokes, Bleses, Basbøll, & Lambertsen, 2012a; Stokes, Kern, & dos Santos, 2012b). The aim of the study is to bring these three themes together when examining the phonological and lexical development of French-speaking children. Specifically, we investigate the influence of phonetic complexity (PhC), PhP, and ND on vocabulary size.

Lexical selection and avoidance

Observational studies support the idea that children select and avoid words on the basis of their PhP capacities. For example, Ferguson and Farwell (1975) reported on a child who had a disproportionately high number of words containing sibilant consonants [s, z, ʃ, ʧ, ʤ] (e.g. *cereal, shoes, cheese, juice, see, eyes* and *sit*). Since input frequency could explain apparent lexical selection patterns, researchers have used experimental paradigms to provide evidence for articulatory effects on word learning. Schwartz and Leonard (1982) showed that children, aged 1;2 to 1;8, learned to produce nonsense words containing sounds that they could produce more easily than nonsense words containing sounds that they could not produce thus confirming the link between phonological experience and lexical acquisition.

Beyond the first word period, authors have found evidence for phonetic effects on lexical selection by examining the phonological characteristics of children's lexicons at different ages according to the MacArthur Communicative Developmental Inventory (MCDI – Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick, & Reill, 1993). Gayraud and Kern (2007) compared the phonological composition of nouns acquired by French-speaking children at 2;0 and at 2;6. In terms of syllable structure, they observed a decrease in the proportion of CV syllables and an increase in the proportion of CVC syllables between the two ages. Words beginning with nasals decreased whereas words beginning with fricatives increased in frequency. Words beginning with bilabials decreased whereas those beginning with alveolars and velars increased. In sum, findings based on French and several other languages (e.g., Cantonese: Fletcher, Chan, Wong, Stokes, Tardif, & Leung, 2004; English: Stoel-Gammon, 1998) show phonetic differences between the vocabularies of younger and older children which match those of production data. Those features which are less frequent in the lexicons of the younger children are those which are acquired later in production.

Other studies have looked at lexical selection in children of the same age, but who vary according to vocabulary size. Kehoe, Chaplin, Mudry, and Friend (2015) observed phonetic selection tendencies in a group of late talkers. One of the ten late talkers did not select any words with initial clusters and four of them did not select any words with final clusters in comparison to their peers with medium or large vocabularies who selected words with clusters. Further information on lexical selection comes from a study by Kern and dos Santos (2016) which examined whether PhC (as defined by the Index of PhC; Jakielski, 2000) could explain variance in vocabulary size in French-speaking children, aged 2;0 to 2;6. They found that it accounted for very little variance in comparison to other variables (e.g., ND, see below). In sum, more research is needed to clarify the role of lexical selection in vocabulary acquisition.

The phonology of late and precocious talkers

There is a wealth of data supporting the association between PhP and vocabulary size. Children who have exceptionally small vocabularies such as late talkers have very limited phonological abilities. Rescorla and Ratner (1996) found that late talkers vocalized less often, had smaller consonantal and vocalic inventories and employed a more restricted set of syllable shapes than their typically developing peers. At the other end of the spectrum, Smith *et al*. (2006) found that lexically precocious two-year olds were superior to their age-matched peers in terms of the number of singleton consonants correct, the percentage of final consonants correct and in their use of phonological processes. The association between vocabulary size and PhP has been observed in a variety of languages including English, Cypriot Greek, and Cantonese (Fletcher *et al*., 2004; Paul & Jennings, 1992; Petinou & Okalidou, 2006) but, to date, there has been little research on the phonology of late talkers in French.

Adult-centred psycholinguistic studies

Stoel-Gammon (2011) contrasted two different approaches to examining the association between lexical and phonological development: child- versus adult-centred approaches. The above-mentioned themes, *lexical selection* and the *phonology of late and precocious talkers*, are child-centred approaches. In adult-centred approaches, researchers have borrowed constructs from language processing in adults to examine the role played by lexical and sub-lexical patterns in the ambient language. We are interested in those studies which have focused on the role of ND in accounting for vocabulary size in children.

Neighbourhood density (ND) refers to the number of phonological neighbours of a word whereby a phonological neighbour is a word that differs from another word by substitution, deletion, or addition of a sound in any word position. Words which contain many phonological neighbours are said to belong to dense neighbourhoods, whereas those which contain few neighbours belong to sparse neighbourhoods. A series of studies by Stokes and colleagues shows that ND accounts for an exceptionally high proportion of variance in the vocabulary size of children acquiring English, French and Danish (Stokes, 2010; Stokes *et al*., 2012a, 2012b). In all of these studies, they coded the mean ND and WF of one-syllable words appearing in two-year-old children's lexicons. The amount of variance accounted for by ND was 39% in Danish-speaking (n=894; age range 2;2-2;6), 47% in English-speaking (n=222; age range 2;0-2;6) and 53% in French-speaking children (n=208; age range

2;0-2;6). In all cases, WF accounted for a small amount of additional variance (English: 14%; French: 9%; Danish: 3%).

One salient finding from Stokes *et al*.'s research is that children with small vocabularies select words with high ND values. Stokes *et al*. (2012b) posit that words from dense neighbourhoods are less taxing on auditory-verbal short term memories than words from sparse neighbourhoods. By virtue of the fact that they share segments with many words they offer a familiar phonetic stream which facilitates word learning. They hypothesize that all children select words with high NDs at the beginning, but children with low vocabularies continue to adopt this strategy for an extended period, thus, impeding later word learning.

Current Study

This study focuses on phonological and lexical associations in French-speaking children, aged 2;5. We seek to confirm Stokes *et al*.'s findings that children with small vocabularies have significantly higher ND values than children with large vocabularies, and that ND accounts for a large percentage of variance in vocabulary size. We extend previous research by including other phonological variables, such as PhC and PhP alongside ND (see Kern & dos Santos, 2016, for PhC).

The first aim is to examine whether children, separated into groups according to vocabulary size, differ in the phonological and lexical characteristics of their lexicons and in their PhP abilities. We predict that children with small vocabularies select words with phonetically simpler forms and with higher ND values than children with large vocabularies. We also predict that they will have inferior PhP skills compared to children with large vocabularies.

The second aim is to examine how much variance in vocabulary size is accounted for by PhC, ND and PhP. We predict that the highest percentage of variance will be accounted for by ND, but given the frequently cited correlation between vocabulary size and PhP, some additional variance will be accounted for by PhP. Given the lack of research on PhC, we make no specific predictions, although the findings of Kern and dos Santos (2016) suggest that PhC does not play a strong role in accounting for vocabulary size.

Method

This study is part of a larger project whose main purpose was to examine the association between early language comprehension and later language and literacy development. The larger study involved testing 65 French-speaking children longitudinally from the ages of 1;4 through to 5;0 years. In this study, we focus on a sub-sample of these children at a single age-range.

Participants

Participants included 40 monolingual French-speaking children, aged 2;5 (+/- 15 days). Children were selected from the larger data-base on the basis of vocabulary size as determined by their percentile scores on the *Inventaire Français du Développement Communicatif* (IFDC) (Kern & Gayraud, 2010) (the French adaptation of the MCDI). Four groups were formed: 1. Late 1 (n=8; 3 girls) were children whose IFDC scores were at or below the 10^{th} percentile (range=40-221); 2. Late 2 (n=9; 6 girls) were children whose IFDC scores were between the 15^{th} and 25^{th} percentile (range=268-353); 3. Middle (n=11; 5 girls) were children whose IFDC scores were between the 40^{th} and 60^{th} percentile (range=372-474); and 4. Precocious (n=12; 6 girls) were children whose IFDC scores exceeded the 90^{th} percentile (range=572-677). All children had normal hearing, were reported to be in good health and were developing normally.

Procedure

Children attended a single session of 60 minutes in the speech laboratory at the University of Geneva in which they received a battery of tests designed to measure executive functions and comprehension and production of vocabulary and morpho-syntax. They also participated in a play session (20 minutes duration) with Fisher Price farm toys while interacting with one of their parents. The play sessions were recorded using a portable digital tape-recorder (Marantz PMD620). The parents completed the IFDC following the session.

Data-coding

A restricted set of the IFDC was coded for PhC and ND (see Stokes *et al*., 2012b). The restricted set included 12 categories of items considered to represent core vocabulary. **All** of the items in the restricted set (n=518) were coded for PhC using the Index of PhC (Jakielski, 2000). A word received a point if it contained: a dorsal consonant (e.g., *camion* [**k**amjɔ̃] ‘truck’), a fricative or liquid (e.g., *avion* [a**v**jɔ̃] ‘plane’; *balle* [ba**l**] ‘ball’), a final consonant (e.g., *balle* [ba**l**] ‘ball’), three-syllables or more (e.g., *animal* [**animal**] ‘animal’), two or more consonants with different places of articulation (PoA) (e.g., *balle* [**b**a**l**] ‘ball’ which has labial and coronal PoAs), a tautosyllabic cluster (e.g., *crayon* [**kʁ**ejɔ̃] ‘pencil’), or a heterosyllabic cluster (e.g., *tracteur* [tʁa**kt**œːʁ] ‘tractor’). Once coding was completed, we determined the mean PhC value for each child. **One-syllable words** of the restricted set of the IFDC were coded for ND (Stokes *et al*., 2012b) using the values generated by the Lexique3 database, a corpus of adult language (New, Brysbaert, Veronis, & Pallier, 2007). The most frequent phonological form was chosen when two word choices (e.g., *beau*/***belle***) were provided. Once coding was completed, a mean ND values was obtained for one-syllable words in each child’s lexicon.

We analyzed children’s spontaneous language samples using Phon, a software program designed for the analysis of phonological data (Rose, MacWhinney, Byrne, Hedlund, Maddocks, O’Brien, & Wareham, 2006). Each child’s language sample was segmented into utterances, glossed, and phonetically transcribed. Three French-speaking students, who had experience in phonetic transcription, performed the analyses. Calculations of PCC were computed automatically for each child using the Query function in Phon. Three participants were re-transcribed by a second transcriber using the Blind Transcription function in Phon. Point-to-point agreement in terms of consonant transcription was high (ranging from 88% to 93%).

Results

Figure 1 shows a box-plot representation of mean PhC for the French-speaking children separated according to vocabulary size. A one-way analysis of variance (ANOVA) revealed a significant group

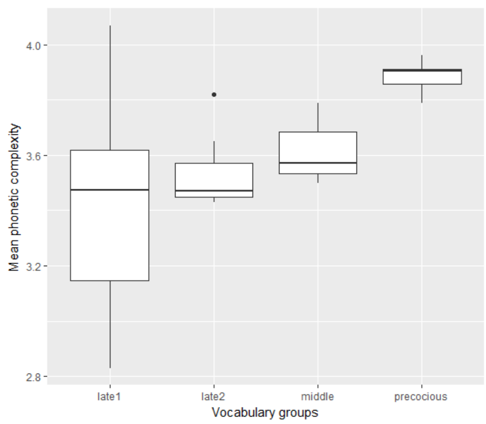


Figure 1. Box plot display of mean PhC for the French-speaking children separated according to vocabulary size

effect (F(3,36)=10.84, p<.001). Children with larger vocabularies had higher PhC values than children with small vocabularies. Tukey HSD multiple comparisons indicated that the precocious group differed significantly from late1 (p<.001), late2 (p<.01), and middle groups (p<.01) but there were no significant differences between the two late and the middle groups.

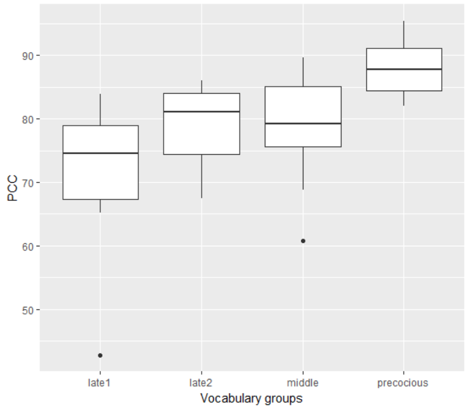


Figure 2. Box plot display of PCCs for the French-speaking children separated according to vocabulary size

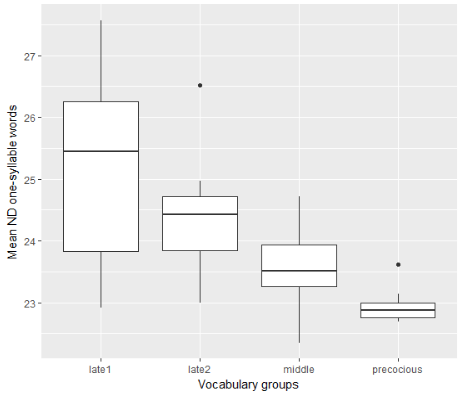


Figure 3. Box plot display of mean ND for the French-speaking children separated according to vocabulary size

Figure 2 provides a box-plot representation of PCCs for the French-speaking children separated according to vocabulary size. A one-way ANOVA revealed a significant group effect ($F(3,36)=7.34$, $p<.001$). Children with larger vocabularies had superior PCCs than children with smaller vocabularies. Tukey HSD multiple comparisons indicated that the precocious group differed significantly from late1 ($p<.001$) and middle groups ($p<.05$) and was marginally different from the late2 group ($p=.06$). There were no significant differences between the other groups.

Figure 3 shows a box-plot representation of mean ND values for one-syllable words in the lexicons of the four groups of French-speaking children. A one-way ANOVA revealed a significant group effect ($F(3,36)=10.93$, $p<.001$). Children with smaller vocabularies had higher ND values than children with larger vocabularies. Tukey HSD multiple comparisons indicated that the precocious group differed significantly from late1 ($p<.001$) and late2 ($p<.01$) but not from the middle group. The middle group differed significantly from the late1 ($p<.01$) but not from the late2 group. The two late groups did not differ significantly from each other.

Finally, we conducted a linear regression entering the three independent variables PhC, ND and PCC as predictors and using vocabulary size at 2;5 as the dependent variable. Together the variables accounted for 71% variance in vocabulary size (adjusted $R^2=.71$). The t values suggested that ND accounted for the most variance in vocabulary size followed by phonetic complexity and PCC. To determine the unique variance of each predictor variable, we entered them in a stepwise fashion starting with ND (see Table 1). ND accounted for 56% unique variance; phonetic complexity, an additional 13% unique variance ($F(1,37)=16.22$, $p<.001$) and PCC, an additional 2% unique variance ($F(1,36)=4.26$, $p<.05$).

Table 1. Coefficients for predicting vocabulary size

Model	B	SE	t	p value
1. Intercept	3064.67	371.41	8.25	p<.001
ND	-111.42	15.52	-7.18	p<.001
2. Intercept	1357.41	527.38	2.57	p<.05
ND	-84.26	14.75	-5.71	p<.001
PhC	290.92	72.22	4.03	p<.001
3. Intercept	1155.38	514.94	2.44	p<.05
ND	-77.7	14.49	-5.36	p<.001
PhC	215.89	78.19	2.76	p<.01
PhP	3.98	1.92	2.06	p<.05

Discussion

The purpose of the study was to examine the influence of lexical selection (measured by PhC), PhP (measured by PCC) and ND on the vocabulary sizes of French-speaking children, aged 2;5. Our results showed significant group differences in all three variables. Children with larger vocabularies selected words with greater PhC and with lower ND values. They had superior PhP abilities compared to children with smaller vocabularies. Multiple comparison tests did not reveal significant differences between all four groups, however. The precocious group, which was characterized by low intra-group variability, differed from the other three groups whereas the late1 group, which was characterized by high intra-group variability, tended to have similar scores to the late2 and middle groups. The lack of significant differences amongst the children with smaller vocabularies could have resulted from reduced power related to sampling effects: the number of items used to determine PhC, ND and PCC was reduced in the late talkers in comparison to the other groups.

Our findings are consistent with those of Stokes and colleagues (2010; 2012b) which show that children with low vocabulary sizes select words from high density neighbourhoods. They studied a group of children ranging in age from 2;0 to 2;6, whereas we focused on children, aged 2;5 only. Thus, our results indicate that even at the outer limits of the age range 2;0 to 2;6, the effects of ND appear to be strong. Our results are also consistent with numerous studies showing that late talkers have inferior PhP abilities and precocious talkers have superior abilities compared to their typically developing peers (Rescorla & Ratner, 1996; Smith *et al*., 2006). As for lexical selection, there has been less research focusing on the PhC of words selected by late and precocious talkers. Our findings, nevertheless, go in the direction of studies which have examined the phonological characteristics of children’s lexicons at different ages (Fletcher *et al*., 2004; Gayraud & Kern, 2007; Stoel-Gammon, 1998). These studies showed that older children’s lexicons contain more phonetically complex words than younger children’s lexicons. Our results indicate that the lexicons of lexically advanced children contain more phonetically complex words than the lexicons of less-advanced children.

Linear regression models revealed that all three variables accounted for a high proportion of variance in vocabulary size (71%). Nevertheless, the bulk of the effect was carried by ND. The value of 56% is not very different from the one reported by Stokes *et al*. (2012b) for French-speaking children (53%) and is higher than the ones reported for English (47%) and Danish-speaking children (39%), suggesting that ND plays a particularly strong role in French in comparison to other languages. The next most important variable was PhC which added 13% unique variance, implying that children with low vocabularies seek phonetically simple words regardless of the density of the neighbourhoods in which these words are found.

Finally, we found that PhP plays a role, albeit a small one, in accounting for children’s vocabulary sizes. Many authors have pondered the nature of the relationship between the two variables: whether having larger vocabularies leads to greater experience in producing sounds or whether greater experience producing sounds leads to increased ability to learn new words (Smith *et al*., 2007). Although the relationship is likely to be reciprocal, our findings support the fact that poor production skills are a limiting factor in the development of a lexicon. In sum, our findings confirm those of previous studies in pointing to a central role of ND in accounting for variance in vocabulary size. They extend research by implicating the role of other phonological variables such as PhC and PhP in lexical acquisition.

Acknowledgements

This study was supported by a grant (#R01HD068458) from the National Institute of Child Health and Human Development awarded to Margaret Friend, Diane Poulin-Dubois and Pascal Zesiger.

References

Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego: CA Singular Publishing Group.

Ferguson, C., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, *51*, 419-439.

Fletcher, P., Chan, C., Wong, P., Stokes, S., Tardif, T. & Leung, S. (2004). The interface between phonetic and lexical abilities in early Cantonese language development. *Clinical Linguistics and Phonetics, 18,* 535-545.

Gayraud, F., & Kern, S. (2007). Caractéristiques phonologiques des noms en fonction de l'âge d'acquisition. *Enfance*, *59*, 324-338.

Jakielski, K. J. (2000, June). *Quantifying phonetic complexity in words: An experimental index*. Paper presented at the Annual Child Phonology Conference. Cedar Falls, IA.

Kehoe, M., Chaplin, E., Mudry, P., & Friend, M. (2015). La relation entre le développement du lexique et de la phonologie chez les enfants francophones. *Rééducation Orthophonique*, *263*, 61-85.

Kern, S., & dos Santos, C. (2016, September). Influence of frequency, neighbourhood density and phonetic complexity on first words. *Paper given at the VIII Congresso Internacional de Adquisición del Lenguaje*. Palma, Spain.

Kern, S., & Gayraud, F. (2010). *L'inventaire français du développement communicatif*. Grenoble: Editions La Cigale.

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, *28*, 661-677.

Paul, R., & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech, Language, and Hearing Research*, *35*, 99-107.

Petinou, K., & Okalidou, A. (2006). Speech patterns in Cypriot-Greek late talkers. *Applied Psycholinguistics*, *27*, 335-353.

Rescorla, L., & Ratner, N. B. (1996). Phonetic profiles of toddlers with specific expressive language impairment (SLI-E). *Journal of Speech, Language, and Hearing Research*, *39*, 153-165.

Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., & Wareham, T. (2006). Introducing Phon : A software solution for the study of phonological acquisition. In D. Bamman, T. Magnitskaia, & C. Zaller (eds.), *Proceedings of the 30th Boston University Conference on Language Development* (pp. 489-500). Somerville, MA: Cascadilla Press.

Schwartz, R., & Leonard, L. (1982). Do children pick and choose? An examination of phonological selection and avoidance in early lexical acquisition. *Journal of Child Language*, *9*, 319-336.

Smith, B. L., McGregor, K. K., & Demille, D. (2006). Phonological development in lexically precocious 2-year-olds. *Applied Psycholinguistics*, *27*, 355-375.

Stoel-Gammon, C. (1998). Sounds and words in early language acquisition. The relationship between lexical and phonological development. In R. Paul (ed.), *Exploring the speech-language connection* (pp. 25-52). Baltimore: Paul Brookes Publishing.

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language*, *38*, 1-34.

Stokes, S. (2010). Neighborhood density and word frequency in toddlers. *Journal of Speech, Language and Hearing Research*, *53*, 670-683.

Stokes, S., Bleses, D., Basbøll, H., & Lambertsen, C. (2012a). Statistical learning in emerging lexicons: The case of Danish. *Journal of Speech, Language, and Hearing Research*, *55*, 1265-1273.

Stokes, S. F., Kern, S., & dos Santos, C. D. (2012b). Extended Statistical Learning as an account for slow vocabulary growth. *Journal of Child Language*, *39*, 105-129.

To cite this article: Kehoe, M. & Patrucco-Nanchen, T. (2017). Phonology of late and precocious talkers in French. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 155-161). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Factors affecting the perception of plosives in second language English by first language Cypriot-Greek listeners

Elena Kkese, Kakia Petinou
elenakkese@hotmail.com, kakia.petinou@cut.ac.cy

Cyprus University of Technology

Abstract. This paper investigates the difficulties adult second language users of English encounter with plosive consonants. It presents the results of a word identification task examining the acquisition of plosive voicing contrasts by college students with Cypriot-Greek background. The task using minimal pair words focused on investigating possible factors affecting plosive identification. Both descriptive and inferential analyses were used for identifying how important each factor is when it comes to plosive consonants. The results provide an indication of the rank order for the examined factors. Specifically, syllable position is identified as having the greatest influence on plosive identification, followed by voicing, word position, and place of articulation for both kinds of analyses. By accepting the hypothesis that less successful differentiation of plosive consonants in the second language on the part of Cypriot-Greek users was partially due to the investigated factors manifested implies that the specific second language sounds do not exist or are non-contrastive in the first language. Nonetheless, because the weighting of auditory cues in the categorization of plosives is language-specific, participants were modifying their identification of voiced plosives to fit the mother tongue. Speech perception can, therefore, account for the data of the present study. Specifically, Voice Onset Time provides important information on the voiceless-voiced distinction as well as word-initial, syllable onset plosive consonants.Taken together, the results of the present study indicate that when dealing with contrastive categories in the second language, the acoustic cue of Voice Onset Time is of crucial importance. For Cypriot-Greek users, acquiring voiced plosives means acquiring new Voice Onset Time patterns.

Keywords: second language, speech perception, plosive consonants, voice onset time

Introduction

Plosive consonants in English and CG

A comparison of the plosive system of English and Cypriot-Greek (CG) indicates that both languages have plosive consonants at three places of articulation: bilabial, alveolar and velar. English and CG plosive consonants, however, differ in the number of plosives and their acoustic realisations. These differences may actually be the reason for the difficulties of second language (L2) users when attempting to acquire the L2 plosive system.

Voice Onset Time (VOT) is normally used to characterise plosive consonants across languages. VOT, the period between the plosive closure release and the beginning of voicing, is the primary acoustic cue for the voicing distinction (Lisker & Abramson, 1964). Three patterns for VOT production are evident (Lisker & Abramson, 1964) involving the long voicing lead in which phonation begins before the oral release (voiced plosives), the short voicing lag in which phonation begins just after the oral release (voiceless unaspirated plosives), and the long voicing lag in which phonation begins after the oral release (voiceless aspirated plosives). In English, plosives can be produced with a long voicing lead and a short voicing lag (voiced *vs.* voiceless) (Okalidou, Petinou, Theodorou, & Karasimou, 2010). Therefore, English is a two-category language consisting of voiced and voiceless plosive consonants. Specifically, VOT is considerably longer for voiceless plosives than voiced ones yielding approximate mean values of 58ms for [p] versus 1ms for [b], 70ms for [t] versus 5ms for [d], and 80ms for [k] versus 21ms for [g] (Lisker & Abramson, 1964).

In CG though, the situation is more complex regarding plosive consonants. The two views proposed for the plosive system agree on the presence of short versus long lag times, but there is disagreement whether voiced plosives are contrastive segments. With reference to VOT, this is considerably longer for voiceless aspirated plosives compared to voiceless unaspirated ones with mean values of 55ms for [pʰ] versus 5ms for [p], 60ms for [tʰ] versus 17ms for [t], and 65ms for [kʰ] versus 22ms for [k] (Arvaniti, 1999; Tserdanelis & Arvaniti, 2001). This suggests that CG voiceless plosives differ from English ones, while voiceless aspirated plosives are the ones that are closer to the English voiceless ones. With reference to voiced plosives, descriptions vary considerably. According to Arvaniti (2010), CG consists of unaspirated and aspirated voiceless plosives, while it has no voiced plosives. This explains why English words such as *league* [liːg] may be pronounced as [lik]. Further descriptions maintain that voiced plosives do exist in CG (Botinis, Christofi, Themistocleous, & Kyprianou, 2004; Okalidou *et al*., 2010). Based on these accounts, plosive consonants can be divided into three voicing categories, namely, voiced unaspirated plosives, voiceless unaspirated plosives, and voiceless aspirated plosives. Regarding voiced unaspirated plosives, they are always found in a post-nasal position as in [kuˈmbin] 'button'. The second and third categories are manifested in the minimal pair [kuˈpin] 'oar' and [kuˈpʰːin] 'small bowl'.

Method

Research question

Through descriptive and inferential analyses, the study attempted to answer the question: What are the factors that affect the identification of plosive consonants for CG users of L2 English? Specifically:

a. What is the effect of consonant voicing in the identification of plosives?
b. How is the identification of plosives influenced by their respective place of articulation?
c. In which word position(s) (word-initial, -medial, -final) are plosives most easily identified?
d. In which position in a syllable are plosives more easily identified (onset/coda)?

The research approach used was quantitative aiming at identifying the factors affecting plosive identification. Differences were examined in the dependent variable (percentage of correctness) thought to be caused by the independent variables (four aforementioned factors).

Word identification task

For the developed task, a total of 120 target items were compiled that were arranged in 60 minimal pairs focusing on the voicing contrast of plosive consonants. Low-frequency words were preferred because they cannot be identified on the basis of fewer perceptual features. Nonetheless, the words had transparent spelling. Each pair of words was parallel in distribution and semantically contrastive differing in only one sound that could be found word-initially, -medially, or -finally (i.e. word-initial: *palate-ballot*, *tessellated-desolated*, *crypt-gripped*; word-medial: *apace-abase*, *mettlesome-meddlesome*, *lacquered-laggard*; word-final: *grippe-grebe*, *alight-allied*, *burke-berg*).

16 minimal pair targets were included for each category of consonants while distractors focusing on the voicing contrast were also intermixed and made up 12 of the minimal pair words. Specifically, 2 to 4 distractors were used for every 16 presentations. These included fricative consonants such as the labiodental [f] and [v], dental [θ] and [ð], and alveolar [s] and [z], while some involved the palato-alveolar affricate consonants [ʧ] and [ʤ]. The words were presented in two fully randomised blocks in order to exclude any systematic patterning while two versions of this task were created, in which the selection of items was entirely complementary. This type of task was chosen to eliminate any semantic information from the input (context-free) as it would have happened if a conversation was presented instead.

Participants and procedure

113 CG users of L2 English with typical speech and hearing were recruited for the purpose of this study. In order to ensure a homogenous participant pool (introductory level students), the participants were first-year students in an English-speaking college. In that way, participants shared the same characteristics (e.g., first language, educational level, socio-economic status) since the aim was to

eliminate inter-group differences. Participation in the task was on a completely voluntary basis and students were ensured about the confidentiality of their personal details. The only cases in which participants were excluded from the sample involved students whose first language (L1) was not CG.

The research period involved three spring semesters in order to investigate whether different students of the same level and background face the same difficulties with the specific sounds. The task was pre-recorded using Audacity 1.3 Beta software for recording and editing sounds. The speaker (one woman, aged 30) was a native speaker of RP (Received Pronunciation). She was told to read at her normal pace without any particular attention to clarity and to imagine that the intended listeners were highly familiar with her voice. The task was administered as a two-alternative forced-choice task via a circling response mode. On each trial, participants listened to a target word along with its foil and responded by circling the word heard. Responses were scored as correct or incorrect generating an overall percent correct score, as well as percent correct scores for the feature classes of consonant voicing, place of articulation, word position, and syllable position. Participants had the opportunity to listen to the words twice while response order and stimulus were counterbalanced.

Results

Quantitative findings: factors affecting plosive identification

Since two different formats of the task were administered, independent sample *t*-tests were conducted to compare mean performance overall. The performance indicated that there was not a significant difference between the two versions of the task, which were administered to the two groups ($p > .05$).

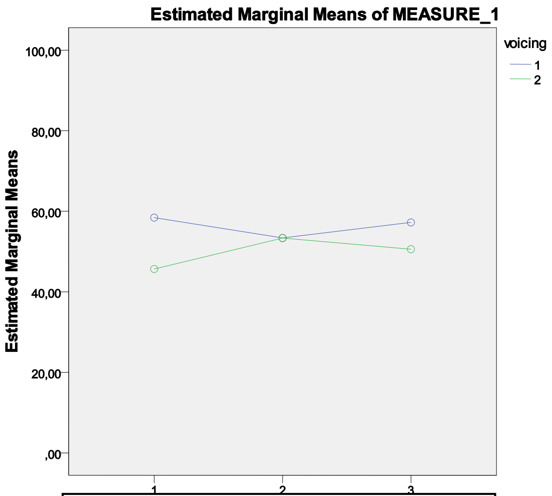
Effect of voicing

A MANOVA (Multivariate Analysis of Variance) with three articulation levels (bilabial, alveolar, velar) and two voicing levels (voiceless, voiced) as within subject factors was conducted in order to compare effects of voicing and place of articulation on performance. The MANOVA indicated a significant main effect of voicing on performance, $F(1,112) = 16.50$, $p < .001$, $\eta^2 = .13$ with overall performance (% correct answers) being higher for voiceless plosive consonants ($M = 56.33$, $SE = 2.15$) than voiced ($M = 49.85$, $SE = 1.90$). There were no significant main effects for place of articulation but the interaction between voicing X place of articulation was significant $F(2,224) = 7.28$, $p < .001$, $\eta^2 = .06$, such that participants performed considerably better in perceiving voiceless plosives only when place of articulation was bilabial (for [p] % correct $M = 58.41$, $SD = 28.66$; for [b] $M = 45.65$, $SD = 23.36$) or velar (for [k] % correct $M = 57.23$, $SD = 25.85$; for [g] $M = 50.57$, $SD = 25.27$), but not alveolar (for [t] % correct $M = 53.36$, $SD = 26.89$; for [d] $M = 53.33$, $SD = 26.03$). Specifically, there was an advantage of bilabial (voiceless $M = 58.41$, $SE = 2.70$; voiced $M = 45.65$, $SE = 2.20$) and velar plosives (voiceless $M = 57.23$, $SE = 2.43$; voiced $M = 50.57$, $SE = 2.38$) compared to alveolar (voiceless $M = 53.36$, $SE = 2.53$; voiced $M = 53.33$, $SE = 2.45$). The interaction is shown in Figure 1.

Effect of word position

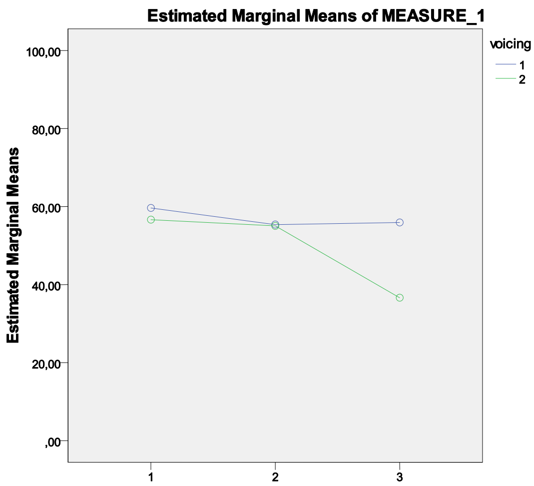
The second MANOVA (voicing X place of articulation X word position) focused on word position as a function of place of articulation and voicing. Significant main effects for voicing (better performance for voiceless), $F(1,112) = 21.60$, $p < .001$, $\eta^2 = .16$ were maintained and an additional main effect for word position was identified, $F(1.8,202.46) = 13.54$, $p < .001$, $\eta^2 = .11$ (Greenhouse-Geisser corrected). For word position, the effect was significant for all three levels with best performance for initial word position ($M = 58.14$, $SE = 2.89$) compared to both medial ($M = 55.22$, $SE = 2.35$) and final ($M = 46.28$, $SE = 1.46$) and with best performance for medial word position compared to final. Especially, the advantage for voiceless ($M = 56.99$, $SE = 2.10$) over voiced plosives ($M = 49.44$, $SE = 1.95$) was greatest for final word position, less pronounced for initial word position, and absent for medial word position. The interaction between word position X voicing was significant $F(2,224) = 18.46$, $p < .001$, $\eta^2 = .14$. Specifically, participants performed significantly better in perceiving voiceless consonants in all three categories (word initial $M = 59.64$, $SE = 3.01$;

word-medial M = 55.39, SE = 2.79; word-final M = 55.92, SE = 2.09) compared to their voiced counterparts (word initial M = 56.64, SE = 3.14; word-medial M = 55.05, SE = 2.55; word-final M = 36.63, SE = 2.01). Contrasts revealed that the percentage of correct responses was significantly higher for medial $F(1,112) = 14.43$ and initial word position $F(1,112) = 19.31$ compared to the final word position, $ps < .001$. However, effect of voicing (advantage for voiceless over voiced) was most pronounced for the word-final category (Figure 2).



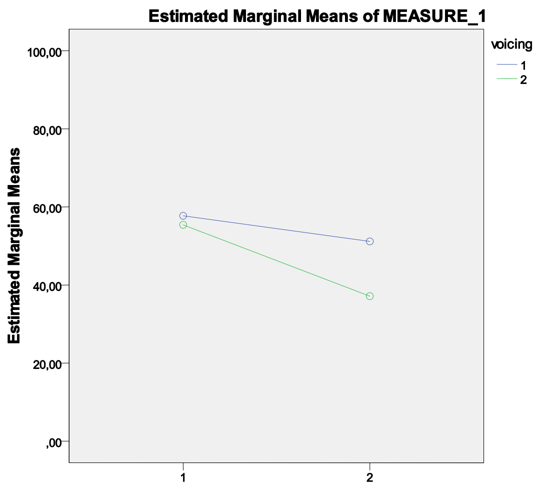
bilabial, alveolar, velar voiceless, voiced

Figure 1. Interaction between voicing X place of articulation (better in bilabial and velar voiceless plosives)



word-initial, word-medial, word-final voiceless, voiced

Figure 2. Interaction between voicing X word syllable position (better in voiceless plosives in word-initial position)



onset, coda voiceless, voiced

Figure 3. Interaction between voicing X position (better in voiceless onset plosives)

Effect of syllable position

The third MANOVA (voicing X place of articulation X syllable) indicated significant main effects for all variables: voicing $F(1,112) = 20.76$, $p < .001$, $\eta^2 = .16$, place of articulation $F(2,224) = 3.66$, $p < .05$, $\eta^2 = .03$, and syllable $F(1,112) = 30.12$, $p < .001$, $\eta^2 = .21$. Concerning syllable position, performance was significantly better for onset (*M* = 56.57, *SE* = 2.46) compared to coda (*M* = 44.15, *SE* = 1.29), especially for bilabial consonants (for [p] % correct *M* = 60.23, *SD* = 33.54; for [b] *M* = 53.83, *SD* = 32.87) compared to alveolar (for [t] % correct *M* = 53.51, *SD* = 32.11; for [d] *M* = 57.20, *SD* = 30.90) and velar consonants (for [k] % correct *M* = 59.42, *SD* = 31.59; for [g] *M* = 55.21, *SD* = 31.82). A significant interaction was identified between syllable X voicing $F(1,112) = 17.33$ $p < .001$, $\eta^2 = .13$, which indicated that participants did better in voiceless consonants especially at onset position (voiceless *M* = 57.72, *SE* = 2.65; voiced *M* = 55.42, *SE* = 2.56) compared to coda (voiceless *M* = 51.16, *SE* = 1.72; voiced *M* = 37.14, *SE* = 2.02). However, effect of voicing (advantage for voiceless over voiced) was most pronounced for coda (Figure 3).

Rank order of the factors affecting plosive consonant identification

Descriptive statistics combining information about all investigated factors regarding the word identification task involving voicing, place of articulation, word position, and syllable position was used to determine frequency of factors. From a descriptive look, by ranking the means for all combinations in descending order (see Table 1), it seems that phonemes in the first (best performance) rows of this table tend to be the ones in onset syllable position.

Specifically, the actual difference between scores tends to be greater as a function of syllable position rather than a function of the other factors. Voicing seems to have the second greatest influence on plosive consonant identification since voiceless ones have higher scores while the ones with the lowest tend to be voiced. Word position follows and is identified as the third most important factor while the last factor identified involves place of articulation. The different means used for analysing the data, thus, indicate to significance above chance for responses even thought the task involved a forced choice task. These indications based on descriptive statistics are further supported with the Multivariate Analyses of Variance (MANOVAs) that have preceded.

Table 1. Descriptive Statistics for Performance of Participants in terms of the Four Investigated Factors in Descending Order

	Rank Order	***M(SD)***	**N**
Variable 1: word-medial onset [p]	1	62.39(40.52)	113
Variable 2: word-initial onset [k]	2	60.47(38.84)	113
Variable 3: word-initial onset [g]	3	60.32(40.78)	113
Variable 4: word-initial onset [p]	4	59.82(42.47)	113
Variable 5: word-initial onset [t]	5	58.63(37.91)	113
Variable 6: word-final coda [k]	6	58.41(34.08)	113
Variable 7: word-medial onset [d]	7	56.93(35.13)	113
Variable 8: word-initial onset [d]	8	56.86(40.12)	113
Variable 9: word-medial onset [b]	9	56.19(39.46)	113
Variable 10: word-final coda [p]	10	54.72(35.18)	113
Variable 11: word-final coda [t]	11	54.65(37.58)	113
Variable 12: word-initial onset [b]	12	52.74(41.06)	113
Variable 13: word-medial onset [k]	12	52.74(42.18)	113
Variable 14: word-medial onset [g]	13	52.04(40.23)	113
Variable 15: word-medial onset [t]	14	51.03(35.79)	113
Variable 16: word-final coda [d]	15	38.50(38.53)	113
Variable 17: word-final coda [g]	16	36.58(34.42)	113
Variable 18: word-final coda [b]	17	34.81(33.45)	113

Discussion and concluding remarks

The results of the present study seem to be straightforwardly related to previous research in language acquisition. Specifically, a number of studies suggested that the less successful differentiation of plosive contrasts was the outcome of the investigated factors. L2 contrasts may often be difficult for users to perceive since they may not be skilled at attending to the needed acoustic cue or set of cues. Further, phonological contrasts in different L1s may be realised differently (acoustically speaking) that results in some degree of L1 influence on the weighting of acoustic cues in perception of the L2 users (depending on their L1 background). On the other hand, L1 users are able to slowly build perceptual categories by being exposed to meaningful input with no interference from an additional language. As a result, the examination of the different factors provides an indication of whether CG users are able to categorise L2 speech sounds into newly formed L2 phonological categories or whether they simply assimilate L2 speech sounds into existing L1 phonological categories. Based on the results, CG users must have perceived voicing, but because the weighting of auditory cues in the categorisation of plosives is language-specific, they were modifying their identification of voiced plosives to fit the L1.

In this context, the acoustic cue of VOT seems to be of crucial importance when dealing with contrastive L2 categories (Arvaniti, 1999; Okalidou *et al*., 2010; Tserdanelis & Arvaniti, 2001; Lisker & Abramson, 1964). VOT, thus, provides important information for the voiceless-voiced distinction since voiceless plosives involve a longer VOT production compared to their voiced counterparts. Concerning place of articulation, velar plosives are produced with longer VOT values while bilabial are associated with the shortest VOT values. Next, plosives in onset position seem to be more easily perceived compared to coda position due to the acoustic cues referring to voicing and place of articulation that may not occur for plosives in coda position. Further, plosives in word initial position are not affected by phonological processes as in other positions.

Lastly, speech perception can also explain why the participants were more successful identifying velar consonants compared to their bilabial and alveolar counterparts, since the former are associated with significantly longer VOT values than the other plosive consonants (Lisker & Abramson, 1964). Even if performance was also good concerning bilabial plosives, participants were generally more successful identifying velar plosives, both voiceless and voiced, while for bilabial plosives the same pattern was not observed since participants were better with voiceless bilabial plosives but not with their voiced counterparts. Since this study concentrated on plosive consonants in L2 English at the word-level, one issue that needs to be addressed in future research is whether a purely phonetic task involving only syllables such as an ABX test or a task involving words in sentences produce different results.

References

Arvaniti, A. (1999). Greek voiced stops: prosody, syllabification, underlying representations or selection of the optimal? *Proceedings of the 3rd International Linguistics Conference for the Greek Language* (pp. 383-390). Athens: Ellinika Grammata.

Botinis, A., Christofi, M., Themistocleous, C., & Kyprianou, A. (2004). Duration correlates of stop consonants in CG *Proceedings, FONETIK 2004*, Department of Linguistics, Stockholm University.

Lisker, L., & Abramson, A.S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, *20*, 384-422.

Okalidou, A., Petinou, K., Theodorou, E., & Karasimou, E. (2010). Development of voice onset time in standard-Greek and Cypriot-Greek-speaking preschoolers. *Clinical Linguistics & Phonetics*, *24*(7), 503-519.

Tserdanelis, G., & Arvaniti, A. (2001). The acoustic characteristics of geminate consonants in CG. *Proceedings of the 4th International Conference on Greek Linguistics* (pp. 29-36). Thessaloniki, Greece: University Studio Press.

To cite this article: Kkese, E., & Petinou, K. (2017). Factors affecting the perception of plosives in second language English by Cypriot-Greek listeners. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 162-167). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Which language sounds good to you? Researching sources of individual differences in processing unfamiliar non-native vowel contrasts

Vita V. Kogan, Joan C. Mora
vitavkogan@gmail.com, mora@ub.edu

University of Barcelona

Abstract. Although recent studies on second language (L2) speech acquisition acknowledge that individual differences play a crucial role in learning outcomes, the ways in which individual differences in the processing of specific phonetic features in unfamiliar sounds interact with L2 speech learning ability for languages that exploit those phonetic features contrastively have yet to be addressed. The interplay between individual differences and specific phonological features results in various learning outcomes. The goal of the current project is to investigate how native listeners with a shared L1 vary in their ability to perceive and produce three non-native vowel contrasts, where each contrast exploits a representative language-specific phonological feature. Our first hypothesis is that individuals' perceptual ability would vary depending on the feature presented in a contrast (e.g., some individuals will demonstrate more accurate perception of a contrast based on feature A, but not B, while others will exhibit the reverse pattern). Our second hypothesis states that individual perceptual skills for specific phonological features will predict the rate of success in the subsequent production of these features. We investigate 30 Spanish monolinguals varying in acoustic and phonological memory capacity. The participants perform an AX rated discrimination task presenting pairs of stimuli that exploit three non-native vowel contrasts (French /i – y/, Portuguese /i – ĩ/, and Russian /i - ɨ/). In addition, participants perform an imitation task based on words and sentences produced in French, Portuguese and Russian. We expect to obtain individual variation in perceptual sensitivity to the different language-specific features in the non-native contrasts examined, occasionally revealing a listener-contrastive feature 'fit' that would be reflected through the presence of a facilitation effect in the production of this feature in the imitation tasks. Having affinity for specific phonological features can give a learner an initial advantage in acquiring the pronunciation of a language where these features are present.

Keywords: individual differences, phonetic aptitude, L2 perception, foreign accent

Introduction

Adults are known for struggling with new speech sounds. The ability to master a foreign phonological system might be the only ability in language learning that is subjected to a critical period (Christiner & Reiterer, 2015). Often, late second language learners exhibit excellent knowledge of vocabulary and grammar, while still not being able to manage L2 pronunciation – a dissociation of capabilities known as the 'Joseph Conrad Phenomenon' (Reiterer, Hu, Erb, Rota, Nardo, Grodd, Winkler, & Ackermann, 2011). Current L2 speech learning models (Best & Taylor, 2007; Escudero, 2005; Flege, 1995) aim to explain the relative difficulty faced by learners in acquiring non-native sounds. The general assumption is that native phonetic categories function as a filter that removes specific acoustic properties of non-native sounds that are not exploited in the L1 and matches the input to the native language environment. This process, called 'equivalence classification', is responsible for inaccurate L2 category formation (Flege, 1995, 2003) and, as a result, causes deficient L2 perception and accented speech.

Despite the powerful influence of a native sound inventory, some individuals still manage to achieve a near-native phonological proficiency (Moyer, 2014). Cases have been reported in second language acquisition research of learners demonstrating exceptional outcomes in phonology even when acquiring a second language later in life (Bongaerts, Planken, & Schils, 1995; Ioup, Boustagi, El Tigi, & Moselle, 1994). Previous research on the role of aptitude in L2 phonological acquisition has shown that a number of cognitive abilities influence L2 phonological performance; the most salient ones being phonological short-term memory (Aliaga-Garcia, Mora, & Cervino-Povedano, 2011) and acoustic memory (Safronova & Mora, 2012a, 2012b). Both types of memory have been found to be related to successful L2 phonological acquisition. Recent studies suggest that greater phonological short-term memory capacity and greater acoustic memory capacity are related to L2 learners' accurate perception of unfamiliar sounds. Among other candidates that play a role in L2 phonological processing and acquisition are: musical ability (Christiner & Reiterer, 2013), attentional control (van Heuven, Conklin, Coderre, Guo, & Dijkstra, 2011), and inhibition (Darcy, Mora, & Daidone, 2016).

The concept of phonetic talent has a long history beginning with Carroll's phonemic coding ability (1981), which was defined as the ability to recognize new sounds and store them in long-term memory. This description of the construct resembles closely the phonological loop, the sub-component of working memory, described by Baddeley (1986) and associated with phonological short-term memory. Both Carroll and Baddeley approach cognitive ability for processing unfamiliar sounds as natural and inherent and free of educational influence. More recent perspectives on phonetic talent (Delvaux, Huet, Piccaluga, & Harmegnies, 2014) focus on the emergent and ever-evolving nature of this ability. According to this view, phonetic aptitude is no longer a fixed innate characteristic of a person, but a skill that emerges from the interaction between speaker-specific psycholinguistic determinants. The researchers that support this conceptualization emphasize the importance of other factors that contribute to mastery of L2 speech, such as articulatory control and the ability to benefit from extensive exposure to a foreign language.

Whereas the conceptualization of aptitude changes over time, what stays the same is the notion that the same set of abilities, innate or developed, is used to acquire phonological features across different languages. In other words, if a foreign language learner is phonetically gifted, his/her gift will be equally beneficial for acquiring labialization in front vowels and vowel nasalization, but he/she might have a more difficult time acquiring L2-specific phonological contrasts based on universally marked phonetic features (e.g., labialization in front vowels) than those based on less marked features (e.g., vowel length). However, that is not what language educators observe in class (MacWhinney, 1995). It is more likely that some phonological features are easier for one student and other features are easier for another student, even if both students share the same mother tongue. A learner-feature idiosyncratic interaction might result in a learning pattern that does not always go along with the developmental trajectory predicted by the existing L2 speech models or current aptitude research. This phenomenon has not yet been addressed by current research.

A reason could be partially due to what Chan, Skehan, and Gong (2011:1) describe as 'the separation of the [linguistic] discipline into two camps': on one hand there are studies that investigate general acquisition processes and analyze group results rather than individual patterns (Mitchell & Miles, 2004); on the other hand, there is a research paradigm that focuses on individual differences without taking into consideration specific L2s (Abrahamsson, & Hyltenstam, 2008; Grey, Williams, & Rebuschat, 2015; Grigorenko, Sternberg, & Ehrman, 2000). While each of the approaches mentioned above shed light on L2 acquisition, both approaches should be taken into consideration to investigate the ways in which language structures interact with individual differences. The current understanding of language as a complex adaptive system (Larsen-Freeman, 2012) inspired L2 research and perspectives that have been interactional in nature (e.g., the relationships between aptitude and motivation, Dörnyei, 2009), but the extent of this theorizing has been rather limited.

Another reason for the lack of research in this topic is an ever-present gap between the research community and language educators that prevent the insights originated in the classroom from moving to the laboratory, where they can be analyzed and explained (to be fair, the opposite is also true: teaching practices often do not take into account current SLA research (Darcy, 2017)). Although linguistics and language teaching seem to interact with each other, their mutual effects upon each other are not equal. The research questions proposed in this article resulted from more than five years of teaching experience in an intensive language program at the Defense Language Institute, USA, where one of the authors worked as a language instructor. During this time, this author recorded and analyzed 120 hours of individual interviews with exceptionally talented students of more than ten languages. The data revealed that some learners have preferences for particular linguistic properties, such as a specific phonological inventory, poverty or richness of inflections, a certain word order, etc. Since all the respondents were native speakers of American English, it was assumed that such preferences are not related to their L1. A number of learners reported an unusual ease with certain combinations of sounds or sound properties that their classmates, in contrast, found particularly challenging. This goes along with Hanulikova, Dediu, Fang, Basnakova, and Huettig's study (2012) that demonstrates striking individual differences in learning Slovak consonant clusters at the very onset of L2 acquisition. Therefore, the purpose of this article is to explore the role of feature-specific aptitude in helping us understand why certain individuals exhibit a propensity to successfully process some phonological features, such as nasalization or linguistic tone, but not others. There are only a few studies suggesting that the interaction between learners' various skill profiles and specific linguistic features result in diverse learning outcomes (MacWhinney, 1995). The practical application of such research is enormous since it would allow for more accurate predictions about individual differences in the learning of specific languages. For example, an individual aptitude for nasalization might facilitate the acquisition of languages where this feature functions contrastively.

Method

Research Questions

The current project addresses the following research questions:

(i) Does an individual's perceptual ability vary as a function of the L0 phonological feature? In other words, do individuals exhibit a talent for a specific non-native phonological feature that is not used in their L1 contrastively?
(ii) If individuals have an aptitude for perceiving specific features, will they also succeed in exploiting these features more effectively in the L2?

Participants

Participants are Spanish functional monolinguals without any prior knowledge of the languages in the study (i.e. Russian, Portuguese, and French). According to the definition provided by Best and Tyler (2007), functional monolinguals are individuals who have been raised in monolingual homes and have not learned another language prior to attending school. In school, they only have basic knowledge of English, i.e. basic classroom instruction and grammar, have not resided in an English speaking country for longer than a month and use only Spanish in their everyday life.

As a part of the experiment, participants complete a questionnaire (adapted from Hanulíková, Dediu, Fang, Bašnaková, & Huettig, 2012) to obtain their demographics and information about their experience with the languages of the experiments, i.e. Russian, Portuguese, and French. This information will be analyzed with respect to the data collected from the other tasks in the study.

Tasks

The proposed data collection procedure comprises three successive parts administered in a single one-hour session. The first part consists of two memory tasks: a serial nonword recognition task and a target sound recognition task. These tasks are designed to evaluate phonological short-term memory and acoustic memory respectively (Baddeley, 1998, 2012) and to tease apart individual differences in memory from individual differences in learner-feature 'fit'. In the first memory task, participants hear 24 pairs of strings of Spanish CVC nonwords increasing in length and are asked whether the order of the nonwords in the sequences is the same or different:

> *Ner* – ***tas*** – ***bol*** – *til*
> *Ner* – ***bol*** – ***tas*** - *til*

The score is obtained by assigning five, six and seven points to the correct responses of five-, six- and seven-item sequences, respectively. In the second memory task, participants listen to a sequence of Spanish sounds and a target sound presented after it and then decide whether the target sound belongs to the sequence or not. In the latter task, the stimuli consist of Spanish CV nonwords, which are manipulated through frequency rotation (speech rotation; Scott, Rosen, Beaman, Davis, & Wise, 2009). This technique preserves the acoustic complexity of the stimuli while making it impossible to be encoded phonologically.

Sequence: *Fe* – ***ro***.
Target sound: ***Ro***.

The second part of the data collection procedure consists of an AX rated discrimination task that is meant to assess the participants' auditory perception of three non-native contrasts: French /i - y/, Portuguese /i – ĩ/, and Russian /i - ɨ/. Each contrast is based on a phonological feature that is representative of a corresponding language (labialization in front vowels, nasalization or front-to-central backness without lip-rounding) and not present in participants' L1. The target vowel contrasts are recorded in the /bVt/ context by female native speakers of French, Portuguese, and Russian. The stimuli are organized into trials where each token A and B is spoken randomly by the same and by a different individual (ISI is 300ms). The participants will be familiarized with all non-native sounds they will hear in the task before the practice block. A-prime score will be calculated for each contrast based on the proportion of 'hits' (the correct selection) and 'false alarms' (the incorrect selection) using Grier's formula (1971), which defines a score of 1.00 as perfect sensitivity and a score of 0.5 as lack of sensitivity.

The last part of the procedure consists of a direct imitation task, during which participants repeat French, Portuguese, and Russian sentences of different lengths and phonetic complexity. This task measures the ability to correctly reproduce speech in an unfamiliar non-native language (L0) produced by a native L0 speaker model, specifically speech focusing on the contrasts of interest (/i - y/, /i – ĩ/, /i - ɨ/). Assessment is performed with online native speaker ratings of participants' speech productions. The raters are instructed to transmit their global intuitive impression of whether the speech sample they are listening to could be spoken by a native speaker of French/Portuguese/Russian (the intuitive rating scale; Jilka, 2009). To ensure the quality of the evaluation procedure, some of the recordings of the native speakers producing the speech samples are randomly inserted into the database. To analyze the data and account for participant-level and stimulus-level variability, a linear mixed effects model is used.

Predictions

We expect to obtain individual variability in perceptual sensitivity to different language-specific features in the non-native contrasts examined, revealing a listener-feature 'fit'. This 'fit' will be further reflected through the presence of a facilitation effect in the production of this feature in the imitation tasks. For example, a number of individuals will exhibit more accurate perception and production of the French contrast /i - y/ while not succeeding with the Russian contrast /i - ɨ/, whereas others will demonstrate a reverse pattern. Some individuals may exhibit aptitude for the Portuguese contrast /i – ĩ/, while failing to discriminate the other two contrasts. By focusing on individual patterns and specific linguistic features and their interaction, we will be able to make a more accurate prediction of L2 development in each learner. From a theoretical perspective, understanding the nature of such learner-feature interactions will shed light on developmental sequences observed in subsequent L2 speech learning. Having an aptitude for specific phonological features can give a learner an initial advantage in acquiring the sound system of a language where these features are used contrastively.

Conclusion

The relationships between individual differences and specific language structures (not only in phonetics but also in other levels of the language, such as morphology and syntax) can produce various learning trajectories. More studies are needed to pursue this insight. While currently we do not have definitive answers on the nature of the proposed relationships, we hope to at least generate productive discussion on this topic and look forward to proposals regarding relevant methodology and data analysis.

References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, *30*(4), 481-509.

Aliaga-Garcia, C., Mora, J. C., & Cerviño-Povedano, E. (2011). Phonological short-term memory and L2 speech learning in adulthood. *Poznan Studies in Contemporary Linguistics*, *47*, 1-14.

Baddeley, A. D. (1986). Working memory. New York, NY: Oxford University Press.

Baddeley, A. D. (1998). Recent developments in working memory. *Current opinion in neurobiology*, *8*(2), 234-238.

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29.

Best, C. T., & Tyler, M. D. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S Bohn (eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Bongaerts, T., Planken, B., & Schils, E. (1995). Can late starters attain a native accent in a foreign language? A test of the critical period hypothesis. In D. Singleton & Z, Lengyel (eds.), *The age factor in second language acquisition* (30-50). Clevedon, UK: Multilingual Matters.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (ed.), *Individual differences and universals in language learning aptitude* (pp. 83-118). Rowley, MA: Newbury House.

Chan, E., Skehan, P., & Gong, G. (2011). Working memory, phonemic coding ability and foreign language aptitude: Potential for construction of specific language aptitude tests–the case of Cantonese. *Ilha do Desterro: A Journal of English Language, Literatures in English and Cultural Studies*, *60*(1), 45-73.

Christiner, M., & Reiterer, S. M. (2013). Song and speech: examining the link between singing talent and speech imitation ability. *Frontiers in Psychology*, *4*:874.

Christiner, M., & Reiterer, S. M. (2015). A Mozart is not a Pavarotti: singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience*, *9*:482.

Darcy, I. (2017). A psycholinguist walks into a classroom: A road-map for bridging research and practice in pronunciation teaching [PowerPoint slides]. Retrieved from https://speechlab.utah.edu/Darcy%20PSLLT%202017%20Keynote.pdf

Darcy, I., Mora, J. C., & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *Language Learning*, *66*(4), 741-773.

Delvaux, V., Huet, K., Piccaluga, M., & Harmegnies, B. (2014). Phonetic compliance: a proof-of-concept study. *Frontiers in Psychology*, *5*:1375.

Dörnyei, Z. (2009). Individual differences: Interplay of learner characteristics and learning environment. *Language Learning*, *59*(s1), 230-248.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Doctoral dissertation. Utrecht, the Netherlands: Utrecht University.

Flege, J. E. (2003). Methods for assessing the perception of vowels in a second language. In E. Fava & A. Mioni (eds.), *Issues in Clinical Linguistics* (pp. 3-28). Padova: UniPress.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Baltimore, MA: York Press.

Grey, S., Williams, J. N., & Rebuschat, P. (2015). Individual differences in incidental language learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences*, *38*, 44-53.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: computing formulas. *Psychological Bulletin*, *75*(6), 424.

Grigorenko, E. L., Sternberg, R. J., & Ehrman, M. (2000). A theory based approach to the measurement of foreign language aptitude: the CANAL-F theory and test. *Modern Language Journal*, *84*, 390-405.

Hanulíková, A., Dediu, D., Fang, Z., Bašnaková, J., & Huettig, F. (2012). Individual differences in the acquisition of a complex L2 phonology: A training study. *Language Learning*, *62*(s2), 79-109.

Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition*, *16*(1), 73-98.

Jilka, M. (2009). Assessment of phonetic ability. *Language Talent and Brain Activity*, *1*, 17-66.

Larsen-Freeman, D. (2012). Complex, dynamic systems: A new transdisciplinary theme for applied linguistics? Language Teaching, *45*(2), 202-214.

MacWhinney, B. (1995). Language-specific prediction in foreign language learning. *Language Testing*, *12*(3), 292-320.

Mitchell, R., & Miles, F. (2004). *Second language learning theories*. London, UK: Hodder Arnold.

Moyer, A. (2014). Exceptional outcomes in L2 phonology: The critical factors of learner engagement and self-regulation. *Applied Linguistics*, *35*(4), 418-440.

Reiterer, S. M., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., ... & Ackermann, H. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: Functional neuro-imaging and brain morphology. *Frontiers in Psychology*, *2*:271.

Safronova, E., & Mora**,** J. C. (2012a) Acoustic and phonological memory in L2 vowel perception. In Martín Alegre, S., Moyer, M., Pladevall, E. & Tubau, S. (eds.), *At a time of crisis: English and American studies in Spain* (pp. 384-390). Department of English and German Philology. Autonomous University of Barcelona/AEDEAN.

Safronova, E., & Mora, J. C. (2012b). Acoustic memory, phonological memory and attention control in L2 speech perception. Paper presented at the *Annual Conference of the American Association for Applied Linguistics* (AAAL), Boston, MA.

Scott, S. K., Rosen, S., Beaman, C. P., Davis, J. P., & Wise, R. J. S. (2009). The neural processing of masked speech: Evidence for different mechanisms in the left and right temporal lobes. *Journal of the Acoustical Society of America*, *125*(3), 1737-1743.

Van Heuven, W. J., Conklin, K., Coderre, E. L., Guo, T., & Dijkstra, T. (2011). The influence of cross-language similarity on within-and between-language Stroop effects in trilinguals. *Frontiers in Psychology*, *2*:374.

To cite this article: Kogan, V. V., & Mora, J. C. (2017). What language sounds good to you? Researching sources of individual differences in processing unfamiliar non-native vowel contrasts. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 168-173). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Anticipatory coarticulation and stability of movements in the fluent speech of Italian children who stutter

Giovanna Lenoci, Irene Ricci
giovanna.lenoci@sns.it, irene.ricci@sns.it

Scuola Normale Superiore

Abstract. This study is part of a project which aims at evaluating speech motor skills during the fluent speech of a cohort of stuttering Italian speakers. Stuttering is a speech motor disorder that may arise from an innate limitation in the speech motor control system, which fails to prepare and organize the movements required for fluent speech (van Lieshout, Hulstijn, & Peters, 2004). Anticipatory coarticulation in CV sequences and stability of speech have been used as a measure of the maturity of articulatory processes in fluent speech production. We want to assess if direct measures of speech dynamics would identify impaired mechanisms in stuttering speech during a phrase repetition task. Ultrasound tongue imaging data for 8 school-aged children, half of them who stutter and the other half who do not stutter, show different articulatory patterns between the two groups, for both motor aspects under investigation. The stuttering group, in fact, shows a higher degree of intra-syllabic coarticulation compared to the control group and decreased stability (i.e., more variability) through multiple repetitions of the same /da/ item. Outcomes of this study suggest that the speech motor control system of children who stutter (henceforth, CWS) is less mature in preparing and executing the speech gestures required for fluent speech. This study contributes to shedding light on the impaired articulatory patterns involved in stuttering speech and to identify the diagnostic markers of the disorder, since they have been evaluated in the speech of children close to the onset of the disorder. Furthermore, ultrasound tongue imaging data are particularly useful in such investigations because they allow deeper and more direct insights into the lingual dynamics of speech production.

Keywords: stuttering, children, ultrasound, coarticulation, variability

Introduction

Stuttering is a speech motor disorder that typically arises in early childhood with unknown causes (Yairi & Ambrose, 2013). It is characterized by disruptions in the flow of speech, also called stuttering-like disfluencies (Yairi & Ambrose, 2005): monosyllabic-word repetitions, word-part repetitions, silent and audible sound prolongations are the hallmark characteristics of the disorder. According to epidemiological estimates, stuttering impacts as many as 5.6% of children and it persists in slightly less than 1% of adults. Despite the high rate of spontaneous recovery (95% of stuttering children) experts claim that identifying children at high risk for persisting stuttering and beginning the therapeutic treatment as soon as possible, increases the chance of total remission (Yairi & Ambrose, 2013). Nevertheless, studying stuttering characteristics in children is important for another reason: to assess the real manifestations of the disorder and to distinguish them from compensatory articulatory strategies that adults can use to overcome stuttering moments. Even if most theories of stuttering postulate the interaction of many factors (i.e., linguistic, cognitive, emotional and genetic variables) in the development and persistence of the disorder, it is clear that the speech motor control system is the critical weak link in the chain of events that lead to speech production (Smith & Kelly, 1997; van Lieshout et al., 2004). According to this perspective, stuttering-like disfluencies are the direct manifestations of an innate limitation of the speech motor system that fails to address the appropriate command signals to drive the muscles involved in speech production. Ultimately, they represent breakdowns in the precisely timed and coordinated articulatory movements required for fluent speech. As a matter of fact, under experimental conditions, people who stutter (henceforth, PWS) perform worse across a range of acoustic measures of speech performance than people who do not stutter (henceforth, PWNS) (Namasivayam & van Lieshout, 2011).

Coarticulation, namely the ubiquitous overlapping of speech gestures, is one of the most relevant aspects of speech motor control and a crucial mechanism for producing fluent speech. It is suggested that the anticipatory movements that we find in anticipatory coarticulation may be disrupted in many types of speech disorder, affecting normal speech motor control such as stuttering (Wingate, 1964). Equivocal results have emerged so far in the literature on coarticulation in stuttering speech, especially because studies vary in terms of the age group of interest, or the measures used to infer coarticulation, and the speech samples which, in the case of stuttering, means that either perceptually fluent or dysfluent tokens were of interest. The present study is specifically devoted to the fluent speech of CWS. The rationale for studying children's speech is that it allows assessment of the real manifestations of the disorder: they are closer to the onset of stuttering compared to adults, and their speech is expected to be free from compensatory articulatory strategies. Studying fluent speech, moreover, allows determining whether the children are speaking abnormally even when they are not stuttering at all. Past research has used acoustic analysis and second formant transitions (F2) for inferences concerning lingual position and movement during speech production. The results of these studies, though, have been quite controversial (Chang, Ohde, & Conture, 2002; Subramanian, Yairi, & Amir, 2003). Interestingly enough, recent studies on speech production with children and adults using ultrasound tongue imaging (henceforth, UTI) have found that direct articulatory measures provide insight into lingual articulation and coarticulation that are not revealed through acoustic measures. UTI is a safe and non-invasive articulatory technique, and it provides clear and valuable information about the shape and the position of the main articulator involved in the production of consonants and vowels, i.e. the tongue. Furthermore, from a clinical point of view, particular attention has been devoted to the potential power of UTI as a biofeedback tool for modifying atypical articulations in speakers with speech disorders (Cleland, Scobbie, & Zharkova, 2016).

While only two UTI studies investigate the coarticulatory patterns in the speech of adults who stutter (Frisch, Maxfield, & Belmont, 2016; Heyde, Scobbie, Lickley, & Drake, 2015), this is the first project devoted to school-aged children. The aim of the study is to evaluate the speech motor skills during a phrase repetition task developed at the ELiTe lab of Scuola Normale Superiore in Pisa. The motor aspects under investigation are those underlying the anticipatory coarticulation and the stability of movements through multiple repetitions of the same item. The latter is another indicator of the speech motor ability: non-pathological developmental studies showed, for example, that a reduction in variability is an indicator of development toward adult-like speech motor control (Zharkova, Lickley & Hardcastle, 2014). Furthermore, PWS were found to be less stable even in their fluent productions as compared to typically fluent speakers. For this reason, we want to investigate this aspect in the speech of CWS, through examining the similarity between tongue postures across multiple repetitions of the same alveolar-vowel sequence. The study was designed to address the following questions: Does the stuttering group show the same stop + vowel coarticulation as control peers? Does the stuttering group show higher variability through multiple repetitions of the same /da/ item?

Method

Participants

Ten school-aged children were employed and recorded for the study: 5 CWS (3 males and 2 females) and 5 CWNS (3 males and 2 females). All participants were monolingual speakers of Italian and they were between 8 and 12 years old. Parents of CWNS reported no history of speech, language and hearing disorders, while parents of the stuttering group confirmed that all children had been formally diagnosed for stuttering by a speech pathologist. Stuttering severity was evaluated using a formal assessment (SSI-3, Riley, 1994) and results classified stuttering severity in a range between mild-to-moderate and moderate-to-severe. None of the parents reported their child to have any neurological, motor, auditory or linguistic impairment that could influence the outcome of the study. Data for 4 CWS and 4 CWNS are presented here.

Stimuli

Target stimuli are combinations of CV syllables with C corresponding to bilabial /b/, alveolar /d/ and velar /g/ stops and V corresponding to high front /i/, low /a/ and high back /u/. The target syllables were embedded in disyllabic pseudo-words of the type /ˈCVba/. Bilabials following the target syllables were used to eliminate the influence of additional lingual coarticulation within pseudo-words. Twelve repetitions of each CV sequence embedded in short carrier phrases (e.g., *la gattina DUba salirà* 'the little cat DUba will go up') were collected in random order, for a total of 108 utterances for each participant.

Procedure

In a child-friendly setup, participants were seated in front of a computer screen in an anechoic chamber in the lab. A micro-convex ultrasound probe (*Mindary probe 65EC10EA*) was held under the chin using the stabilization headset (*Articulate stabilisation headset,* Articulate Instruments ltd) and was oriented to display a mid-sagittal configuration of the tongue. The ultrasound images from a *Mindray UTI system* were acquired at a frequency rate of 60 MHz. Pre-recorded auditory stimuli were played during the recording session and children were asked to repeat exactly what they had heard. Ultrasound data were recorded and analysed by *Articulated Assistant Advanced* software (AAA).

Analysis

First, all fluent CV syllables were phonetically segmented and labelled, using Praat (Boersma & Weenink, 2001). Three intervals were selected on the spectrogram of each CV sequence: 1) the consonantal closure (i.e., from the offset of the preceding vowel to the burst of the stop); 2) the VOT of the C (i.e., from the burst to the first glottal pulse of the following vowel); 3) the V-target (i.e., from the first glottal pulse to the offset of the vowel) (Figure 1). Annotations were imported into AAA and a semi-automatic tongue contour splining was performed in the acoustic interval spanning from the beginning of the consonantal closure to the end of the following vowel.

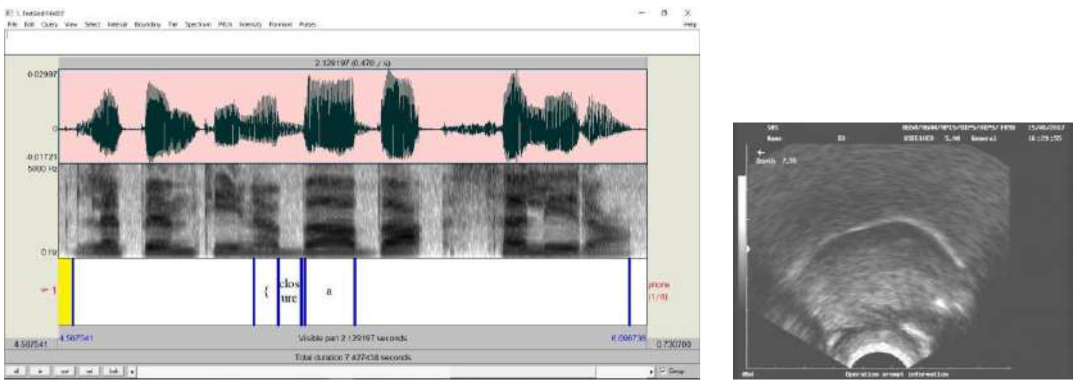


Figure 1. Example of a PRAAT window on the left, with the waveform, the spectrogram and the text-grid containing the segmentation and labelling of a [ba] sequence. On the right, an ultrasound frame at the middle of the consonant /b/ from the word 'baba' produced by a child. The line of brightness in the ultrasound image represents the midsagittal tongue profile

The time points relevant to the analysis presented here were: the temporal midpoint of the consonantal closure and the temporal midpoint of the vowel for the analysis of anticipatory coarticulation; the midpoint of the consonantal closure, for /da/ tokens, for the analysis of token-to-token variation.

For each relevant contour, the x-coordinate of the highest point of the tongue body was automatically extracted and used for subsequent coarticulation analyses. Each x-y coordinate of the tongue contours was used instead for the analysis of token-to-token variability.

Results

In line with previous acoustic studies, we first investigated anticipatory coarticulation using Locus Equations (henceforth, LE). Sussman and collaborators computed linear regressions for the second formant (F2) between the vowel onset and its midpoint to test for the linear relationship between stops and vowels in CV sequences (Sussman, Hoemeke, & McCaffrey, 1992). They found that the slope of the regressions varied with the amount of CV coarticulation: the steeper the slope, the higher the degree of anticipatory coarticulation.

To achieve our purposes, we transposed LE to the articulatory domain and used the horizontal position of the highest point of the tongue instead of F2, in line with previous recent studies which have demonstrated the articulatory basis of LE (Noiray, Menard, & Iskarous, 2013). For regression analyses, we used the highest point of the tongue at the mid-point of the consonantal closure (dependent variable) and at the mid-point of the following vowel (independent variable). LE slope means were calculated separately for the stuttering and the control group, for each place of articulation. Figure 2 displays the resulting regression lines for the bilabial, alveolar and velar places of articulation.

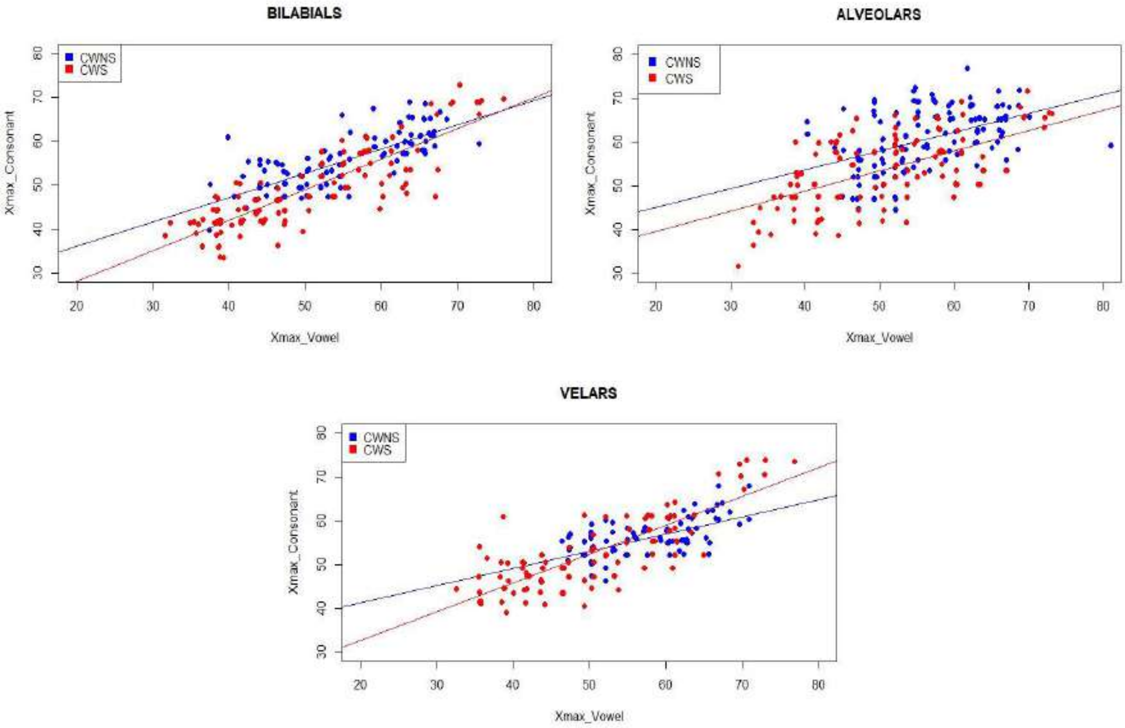


Figure 2. Linear regressions between the consonant closure midpoint and vowel midpoint for the bilabial, alveolar and velar places of articulation. The red lines are the regression functions for CWS, the blue ones are for the CWNS. Slope values are: [bilabial α= 0.694 for the stuttering group and α=0.553 for the control group]; [alveolar α= 0.458 for the stuttering group and α=0.429 for the control group]; [velar α= 0.658 for the stuttering group and α=0.392 for the control group]

Results show that CWS and CWNS exhibited different patterns of coarticulation magnitude according to the place of articulation of the stop, because they vary in the order of slope amplitude: /b/>/g/>/d/ for CWS and /b/>/d/>/g/ for CWNS.

Results also show that, for the three places of articulation, the stuttering group exhibited higher slope values compared to the control group. This means that CWS had a higher degree of gestural overlap during fluent speech compared to the control group.

For speech stability, nearest neighbour distances (henceforth, NND) (Zharkova & Hewlett, 2009) were obtained for tokens with the same vowel context (i.e., /da/ in the analyses presented here). This measurement consists in calculating the mean of all the Euclidean distances between each point on one curve and its nearest neighbour on the other. This average distance between curves within the same vowel context can be interpreted as a measure of speech motor stability for individual speakers. It was determined within each individual and then used in statistical analyses of differences between groups. Results reveal that the stuttering group showed higher level of variability (i.e., less accuracy and stability) for token-to-token variation, compared to the control group. Indeed, a *t*-test showed that NND measure was significantly higher for the first group compared to the second one, as shown in Table 1.

Table 1. NND mean values for the two groups. Significance level 0.05

Group	*N*	*Mean*	*S.d.*
CWS	176	**2.82**	.07
CWNS	176	**2.51**	.08

Discussion

In the present study, ultrasound was used to investigate anticipatory stop-vowel coarticulation in the productions of 8 children, four stuttering and four control speakers. Previous literature has not addressed speech motor stability and lingual coarticulation using articulatory measures in typical and stuttering children, especially in Italian. Some general conclusions can be drawn about articulatory dynamics in the speech of stuttering children. The first one is that the fluent productions of CV syllables from 4 stuttering children and 4 control speakers differ in two aspects of speech motor control, namely coarticulation and stability of speech. A specific trend emerges from the analysis of the anticipatory coarticulation: it was stronger for the stuttering group compared to the control group. This means that CWS reach fluency in speech through a higher spatio-temporal overlap of the speech gestures, which would be a consequence of a more global and undifferentiated planning and programming of the syllables (Nittrouer, Studdert-Kennedy, & Neely, 1996). These results may suggest that in CWS's speech movements of individual articulators over a syllable are more 'global' than in normally fluent peers, resulting in syllable-sized units of production, as opposed to the more dynamic approach taken by CWNS, which affords individual segments more autonomy. Therefore, CWS may have difficulty with transitions and/or blending across sounds resulting in undershooting their articulatory targets.

Significant differences emerged between the two groups, also for the measure of token-to-token variation: CWS were found to be more variable in the accuracy of movements for the same articulatory target, compared to the control group. According to Max, Guenther, Gracco, Glosh and Wallace (2004) an important aspect of stuttering may lie in children's inability to acquire stable and correct mappings between motor commands and sensory consequences. The high instability found in the articulatory patterns of CWS may be a physical manifestation of unstable or insufficiently-activated gestural model of speech production 'to appropriately update these mappings during speech development, or to sufficiently activate and successfully use these mappings for efficient sensorimotor control of speech mechanism' (Max *et al*., 2004:113).

Conclusions

These preliminary results corroborate some previous findings according to which PWS 'may be located more toward the unskilled end of a presumed (normal) speech motor skill continuum' (van Lieshout *et al*., 2004). Even though preliminary, in terms of the low number of subjects, these results allow us to gain a better insight in the articulatory dynamics of children close to the onset of the disorder, and consequently to assess the real manifestations of stuttering. Besides this, UTI data could be very useful from a clinical point of view, because it can be used as a biofeedback tool for modifying atypical articulation in speakers with speech disorders. To this respect, several studies have demonstrated that ultrasound visual biofeedback can be used for demonstrating and treating articulations which are hard to see (Cleland, Scobbie, & Wrench, 2015). Hence, we think that clinicians could use this tool also with CWS to develop therapeutic protocols apt to improve 'motor' knowledge about the correct posture of speech sounds and the transitioning between them.

References

Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*, 341-345.

Chang, S. E., Ohde, R. N., & Conture, E. G. (2002). Coarticulation and formant transition rate in young children who stutter. *Journal of Speech, Language and Hearing Research*, *45*, 676-688.

Cleland, J., Scobbie, J., & Wrench, A. (2015). Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics and Phonetics*, *29*, 575-597.

Cleland, J., Scobbie, J., & Zharkova, N. (2016). Insights from ultrasound: Enhancing our understanding of clinical phonetics. *Clinical Linguistics and Phonetics*, *30*, 171-175.

Frisch, S. A., Maxfield, N., & Belmont, A. (2016). Anticipatory coarticulation and stability of speech in typically fluent speakers and people who stutter. *Clinical Linguistics and Phonetics*, *30*, 277-291.

Heyde, C. J., Scobbie, L. M., Lickley, P., & Drake, E. K. E. (2015). How fluent is the fluent speech of people who stutter? A new approach to measuring kinematics with ultrasound. *Clinical Linguistics and Phonetics*, *30*, 292-312.

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., & Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: A theoretical model of stuttering. *Contemporary Issues in Communication Science and Disorders*, *31*, 105-122.

Namasivayam, A. K., & van Lieshout, P. (2011). Speech motor skill and stuttering. *Journal of Motor Behavior*, *43*, 477-489.

Nittrouer, S., Studdert-Kennedy, M., & Neely, S.T. (1996). How children learn to organize their speech gestures: Further evidence from fricative-vowel syllables. *Journal of Speech and Hearing Research*, *39*, 379-389.

Noiray, A., Menard, L., & Iskarous, K. (2013). The development of motor synergies in children: Ultrasound and acoustic measurements. *Journal of Acoustical Society of America*, *133*, 444-452.

Riley, G. D. (1994). *Stuttering severity instrument for children and adults-3 (SSI-3)*. Austin, TX: Pro-Ed.

Smith, A., & Kelly, E. (1997). Stuttering: A dynamic multifactorial model. In R. Curlee & G. M. Siegel (eds.), *Nature and treatment of stuttering: New directions* (pp. 204-217). Needham Heigts, MA: Allyn & Bacon.

Subramanian, A., Yairi, E., & Amir, O. (2003). Second formant transitions in fluent speech of persistent and recovered preschool children who stutter. *Journal of Communication Disorders*, *36*, 59-75.

Sussman H. M., Hoemeke K. A., & McCaffrey, H. A. (1992). Locus equations as an index of coarticulation for place of articulation distinctions in children. *Journal of Speech and Hearing Research*, *35*, 769-781.

van Lieshout, P. H. H. M., Hulstijn, W., & Peters, H. F. M. (2004). Searching for the weak link in the speech production chain of people who stutters: A motor skill approach. In B. Maassen, R. Kent, H. F. M. Peters, P. H. H. M van Lieshout, & W. Hulstijn (eds.), *Speech motor control in normal and disordered speech* (pp. 313-355). Oxford, UK: Oxford University Press.

Wingate, M. E. (1964). A standard definition of stuttering. *Journal of Speech and Hearing Disorders*, *29*, 484-489.

Yairi, E., & Ambrose, N. (2005). *Early childhood stuttering: For clinicians by clinicians*. Austin, TX: Pro-Ed.

Yairi, E., & Ambrose, N. (2013). Epidemiology of stuttering: 21st century advances. *Journal of Fluency Disorders*, *38*(2), 66-87.

Zharkova, N., & Hewlett, N. 2009. Measuring lingual coarticulation from midsagittal tongue contours: Description and example calculations using English /t/ and /ɑ/. *Journal of Phonetics*, *37*, 248-265.

Zharkova, N., Lickley, R. J., & Hardcastle, W. J. (2014). Development of lingual coarticulation and articulatory constraints between childhood and adolescence: An ultrasound study. In S. Fuchs, M. Grice, A. Hermes, L. Lancia, & D. Mücke (eds.), *Proceedings of the 10th International Seminar on Speech Production* (ISSP), (pp. 472-475). Retrieved from http://www.issp2014.uni-koeln.de/wp-content/uploads/2014/Proceedings_ISSP_revised.pdf

To cite this article: Lenoci, G., & Ricci, I. (2017). Anticipatory articulation and stability of movements in the fluent speech of Italian children who stutter. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 174-179). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Production of English /iː/ and /ɪ/ by Polish learners of English: A longitudinal study

Dorota Lipińska
dorota.lipinska@yahoo.com

University of Silesia

Abstract. Research on speech is usually focused on one of two levels – segmental or suprasegmental. The analysis of vowels belongs to the former one. Foreign vowels usually pose significant difficulties for language learners. It is true not only for English (see studies mentioned in the paragraph below), but also for other languages such as French (e.g., Kaczyński, 1992; Kotuła, 2015). Vowels found in sound systems of various languages not only differ according to their quality (expressed in terms of frontness/backness and openness/closeness) and duration, but they may also be influenced by surrounding consonants to a considerable extent. Moreover, a particular vowel which is theoretically the same phoneme across various accents of one language, may be realised differently leading to some kind of confusion in language learners. The contrast between English /iː/ and /ɪ/ has already been researched and analyzed in quite a few studies, with a focus on both speech production and speech perception (e.g., Bayonas, 2008; Bogacka, 2004; Casillas, 2015; Cebrian, 2006; Lengeris, 2009; Lipińska, 2014; Nowacka, 2010; Rojczyk, 2010; Wong, 2015). However, the number of longitudinal studies based on acoustic analysis is still severely limited. The aim of this study is to analyze the production of the aforementioned vowels by advanced Polish learners of English. 57 students of English philology participated in the study. They were recorded twice (before and after a term of segmental pronunciation course devoted to English vowels) in a sound-proof booth of the Acoustic Laboratory at the Unversity of Silesia, Poland. The analyzed sounds were embedded in carrier words /bVt/ which were in turn embedded in carrier sentences *I say /bVt/ now*. The subjects performed two tasks each time – sentence reading and shadowing after a native speaker of Standard Southern British English. The analysis of results was done with the use of Praat (Boersma, 2001) and revealed that, although at the beginning of the study participants did not differentiate between the two vowels, their pronunciation improved to a considerable extent after the course in English phonetics.

Keywords: L2 pronunciation, vowels, SLA, phonetics

Introduction

When researchers examine non-native speech production and perception, most often they base their assumptions on two models – the Speech Learning Model (SLM) by Flege (1995) and the Perceptual Assimilation Model (PAM) by Best (1995), developed by Best and Tyler (2007). Practically, both of those models suggest that the phonetic similitudes and differences between segments in learners' L1 and L2 are bound to affect the degree of potential success in perceiving and producing non-native sounds. Thus, it can be said that in this case phonetic similarity or dissimilarity will be defined in terms of the articulatory and acoustic characteristics of particular speech sounds (e.g., Rojczyk, 2009, 2010, 2011). Even though there are some studies which test the SLM against the PAM (e.g., Rohena-Madrazo, 2013), both models may be used as complimentary to each other (Best & Tyler, 2007).

The amount of research on non-native vowel production and perception is quite vast. What is particularly interesting for the purpose of this paper, are the studies focusing on English /iː/ vs. /ɪ/ contrast, complemented with the analysis of 'similar' sounds present in other languages. Here one can also find an abundance of research and papers (e.g., Bayonas, 2008; Bogacka, 2004; Casillas, 2015; Cebrian, 2006; Nowacka, 2010; Rojczyk, 2010; Lipińska, 2014; Wong, 2015). These articles have shown that those language learners whose L1s are characterized by more limited vowel systems than English, in which there is no tense/lax vowel contrast, encountered serious difficulties in differentiating between English /iː/ and /ɪ/, and usually treated those vowels as a single category. What is more, they frequently assimilated them to their native 'similar' sound categories (all the findings were in accordance with the assumptions presented in both SLM and PAM). However, having been subjected to appropriate phonetic training, most study participants improved both their vowel production and perception and were able to form new sound categories. In most cases the researchers observed not only some degree of separation of the two non-native sounds, but also the dissociation of foreign and native vowels. Still, sometimes there was no qualitative difference, only quantitative (tense vowels were significantly longer than their lax counterparts). Nevertheless, those studies are unfortunately characterized by quite a few serious drawbacks. Some of them concentrated only one group of EFL/ESL learners at some particular point of their (inter)language development, others were cross-sectional and analyzed the results obtained by various groups of subjects, another group of studies involved only the auditory analysis of the vowels produced by the informants which is by nature very subjective and immeasurable, and the remaining studies included very short phonetic training sessions (e.g., very few pronunciation classes only and the immediate post-test). This is why it would be necessary to design a new, longitudinal study which incorporated the same group of subjects, as well as a long-term phonetic training between the pre-test and the post-test.

Current study

The aim of this paper is to investigate whether advanced L2 learners are able to separate new non-native vowel categories (English /iː/ and /ɪ/) from each other and from their native, neighboring vowels (Polish /i/ and /ɨ/), as well as to what extent an academic course in practical English phonetics and phonology can facilitate this complex process.

Material

The vowels analysed in this study included English /iː/ and /ɪ/, as well as Polish /i/ and /ɨ/. Since the Polish vowel system is very scarce and does not contain any tense/lax vowel contrasts (e.g., Jassem, 2003), it is very difficult for Polish learners of English to form two new various vowel categories for English /iː/ and /ɪ/, and they are usually subsumed into two already existing native sound categories – Polish /i/ and /ɨ/. While at the beginning of ESL/EFL learning both English sounds are mostly realized as Polish /i/, at some later stages, when learners notice that there are some differences between those foreign sounds, English /ɪ/ frequently becomes Polish /ɨ/-like (e.g., Lipińska, 2014; Sobkowiak, 2004). The correct relation of those English and Polish vowels to each other is shown in Figure 1 below.

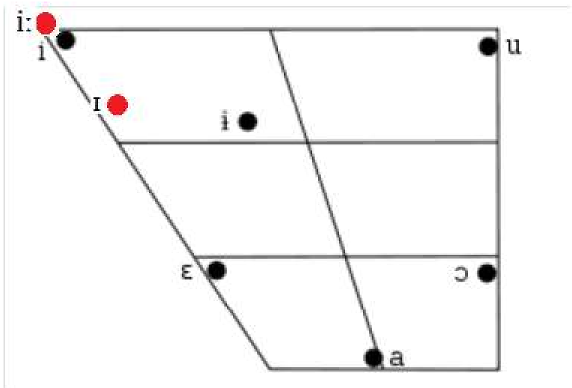


Figure 1. A Polish vowel chart (after Jassem, 2003) with the two English vowels added (red dots; modifications mine)

Study participants

57 subjects, recruited among first-year students of English philology at the Institute of English and the Institute of English Cultures and Literatures (University of Silesia, Poland), participated in the study. They were around 19 years old (mean: 19.4). The administration of a standardized placement test in general English as well as the results of the high-school final exams (so-called Matura) obtained by

the participants allowed the researchers to state that they were upper intermediate/advanced users of English (level B2/C1 according to the Common European Framework of Reference for Languages). Most of them were females (80%), which is rather typical of departments of modern languages. Nobody reported any hearing or speaking disorders. Moreover, all the subjects were volunteers and were not paid for their participation in the study. The study participants were divided into an experimental group and a control group of almost equal size (30 and 27 members, respectively).

Methods

The analysed Polish and English vowels were embedded in a /bVt/ context. In most studies the /hVd/ context is chosen as it has been found to produce the least expansive formant transitions from neighbouring consonants (Peterson & Barney, 1952). However, it was impossible for the Polish-English pair of languages since while English uses a glottal fricative /h/, Polish has a velar /x/ (Jassem, 2003). Subsequently, the /bVt/ words were embedded in carrier sentences *I say /bVt/ now* in English and *Mówię /bVt/ teraz* in Polish. The sentences were presented to the subjects in two forms – an orthographic one and an audio one. One of the tasks involved reading the sentences from the computer screen (an orthographic version of each sentence was displayed together with the IPA transcription of the target word). Another task consisted in semi-shadowing after a native speaker of Standard Southern British English.

There were two recording sessions – the pre-test took place before the beginning of the practical pronunciation course and the post-test was completed after the first part of such a university course (this part focused on segments, vowels for the experimental group, consonants for the control group). The phonetic training between the two recording sessions was very intensive. It lasted one semester (30 hours) and included both perception and production training with the use of various methods and tools. The recording sessions took place at the Acoustic Laboratory of the Institute of English, in a sound-proof booth. The gathered data was subjected to acoustic analysis using Praat (Boersma, 2001). The first two formants were measured at a vowel midpoint and plotted on a vowel plane (after the process of normalization with the use of Lobanov transform). Then in order to identify separate sound categories and to observe how they interact with each other, a Principal Component Analysis (PCA) was used.

The following hypotheses were put forward:

- Hypothesis 1: Prior to the phonetic training English /iː/ will be completely assimilated to Polish /i/
- Hypothesis 2: Prior to the phonetic training English /ɪ/ will be completely assimilated to Polish /i/
- Hypothesis 3: After the phonetic training English /ɪ/ will move towards Polish /ɨ/
- Hypothesis 4: After the phonetic training the English vowels will separate from Polish vowel categories

Results: Pre-test

The results of the pre-test are shown in Figure 2 and Figure 3 below.

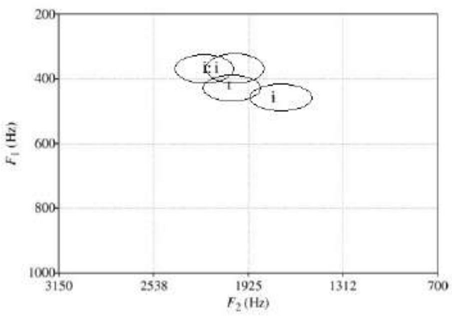


Figure 2. The results obtained by the control group in the pre-test

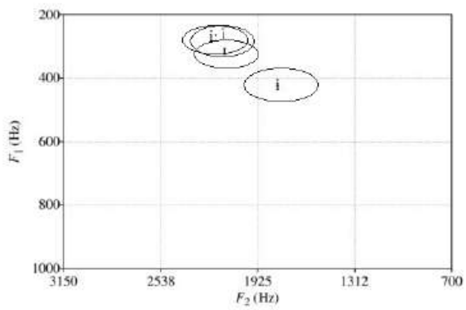


Figure 3. The results obtained by the experimental group in the pre-test

It can be noticed that both members of the control group and the experimental group did not notice any differences between English /iː/ and /ɪ/, and treated them as a single category. Moreover, they replaced the English vowels with one Polish sound: /i/. The sound categories merged completely. The computed Euclidean distance between Polish /i/ and both English sounds was <50 Hz in each case which confirms the observations. At the same time, Polish /ɨ/ remained separated and virtually did not approach any of the foreign vowels (the computed Euclidean distance was >500 Hz). The findings confirmed Hypothesis 1 and Hypothesis 2.

Results: Post-test

The results of the post-test are shown in Figure 4 and Figure 5 below.

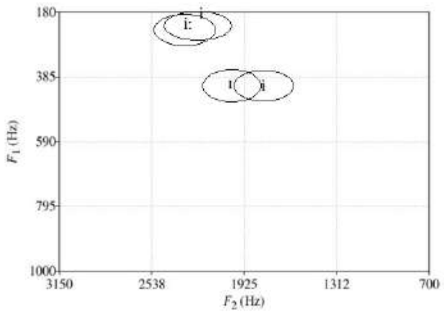


Figure 4. The results obtained by the control group in the post-test

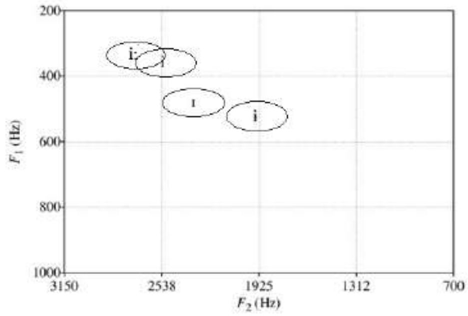


Figure 5. The results obtained by the experimental group in the post-test

As one can easily observe, the intensive training in English phonetics and phonology helped the subjects separate the native and non-native vowel categories. The experimental group managed to form new distinct categories for English /ɪ/ and /iː/ by moving English /ɪ/ towards Polish /ɨ/, as well as find the disparate acoustic space for native and foreign sounds. Still, English /iː / and Polish /i/ shared some of this space, but this is normal since the two sounds are very close to one another (see Figure 1). This fully confirmed Hypothesis 3 and, partially, Hypothesis 4. What is more, the computed Euclidean distance between particular sounds was sufficient to state that the vowel categories separated successfully (about 395 Hz between Polish /ɨ/ and English /ɪ/, about 220 Hz between Polish /i/ and English /ɪ/, about 190 Hz between Polish /i/ and English /iː/ and finally about 400 Hz between English /iː/ and English /ɪ/). The distances between particular sounds and the placement of the analysed vowels in the vowel space allow us to state that the phonemes in question have been separated correctly and the new sound categories have been established successfully.

What is striking is the fact that not only the experimental group improved their performance but, to some extent, the control group also re-formed their sound categories. Even though the subjects from this group were trained in consonants, they noticed that English /ɪ/ differed significantly from English /iː/ and Polish /i/ and hence created a separate category for this sound (partially merged with Polish /ɨ/). The mean distance between Polish /ɨ/ and English /ɪ/ was about 180 Hz while the distance between the close sounds (Polish /i/ and English /iː/) and the near-close phonemes (Polish /ɨ/ and English /ɪ/) was usually about 250 Hz. At the same time Polish /i/ and English /iː/ remained practically the same (the average Euclidean distance < 60 Hz), but the general trend suggests that the phonetic training and drawing students' attention to correct pronunciation has a kind of a side effect – improving not only the elements being trained at a particular moment, but also other elements of speech.

Discussion

The results obtained in this study and described above supplement previous studies on this topic with new, longitudinal data and provide acoustic analysis of the development of new vowel categories by

Polish learners of English. They show that appropriate training in English phonetics and phonology may help separate effectively, not only the non-native sounds from each other but the native and foreign categories, as well. The correct separation of vowels from one another, together with the appropriate phoneme placement in the vowel space, suggests that the analysed sound categories are established optimally. Moreover, as most researchers suggest, the focus on segmental phonetics is of great importance as it is on the particular sounds where mispronunciations and misunderstandings begin.

Conclusion

To conclude, it can be stated that although it may be difficult, Polish learners of English are able to separate English /iː/ and /ɪ/, and Polish /i/ and /ɨ/ categories. Formal instruction in English phonetics and phonology is very helpful in this process and enables language learners to form new vowel categories in their interlanguage. Prior to phonetic training, even though the subjects had already been quite advanced users of English, they had not noticed the sound differences and hence had not been able to speak English correctly. After the first semester of the pronunciation course, ESL/EFL learners trained in English vowels achieved much better results in the production experiment than the subjects from the control group. This proves the effectiveness of such university courses.

Naturally, the study may still be improved and extended. First of all, other vowels and their combinations need to be analyzed. Secondly, it is advisable to carry out a longer longitudinal study (lasting 2, 3 semesters or even more) in order to check the long-term effects of pronunciation training. What is more, it is worth analyzing the same way other pairs of languages (both genetically related and unrelated ones).

References

Bayonas, M. G. (2008). Perception of English vowels as first and second language. *Revista Electrónica de Lingüística Aplicada*, *7*, 79-89.

Best, C. T. (1995). A Direct Realist View of Cross-Language Speech Perception. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. pp. 171-204). Baltimore, MA: York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second language speech perception: Commonalities and complementarities. In O.S-. Bohn & M. J. Munro, M.J. (eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13-34). Amsterdam: John Benjamins.

Boersma, P. (2001). Praat: A system for doing phonetics by computer. *Glot International*, *10*, 341-345.

Bogacka, A. (2004). On the perception of English high vowels by Polish learners of English. In E. Daskalaki, N. Katsos, M. Mavrogiorgos, & M. Reeve (eds.), *CamLing 2004: Proceedings of the University of Cambridge Second Postgraduate Conference in Language Research. Cambridge Institute of Language Research* (pp. 43-50). Cambridge, UK: Cambridge University Press.

Casillas, J. V. (2015). Production and perception of the /i/-/ɪ/ vowel contrast: The case of L2-dominant early learners of English. *Phonetica*, *72*(2-3), 182-205.

Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, *34*, 372-387.

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. pp. 220-273). Timonium: York Press.

Jassem, W. (2003). Illustrations of the IPA: Polish. *Journal of the International Phonetic Association*, *33*(1), 103-107.

Kaczyński, M. (1992). *Materiały pomocnicze do nauczania wymowy francuskiej.* Lublin: Wydawnictwo Uniwersytetu Marii Curie Skłodowskiej.

Kotuła, K. (2015). Nowe media wspomagające autonomię uczniów szkół ponadgimnazjalnych w nauce wymowy języka francuskiego. In M. Pawlak (ed.), *Autonomia w nauce języka obcego – Techniki, strategie, narzędzia.* Konin: Wydawnictwo Państwowej Wyższej Szkoły Zawodowej w Koninie.

Lengeris, A. (2009). Perceptual assimilation and L2 learning: Evidence from the perception of Southern British English vowels by native speakers of Greek and Japanese. *Phoentica*, 66(3), 169-187.

Lipińska, D. (2014). Extension as a simplification strategy: The case of Polish learners of English applying Polish /i/ Category to English /ɪ/ and /iː/ Contexts. In G. Drożdż & A. Łyda, A. (eds.), *Extension and its limits.* Newscastle upon Tyne, UK: Cambridge Scholars Publishing.

Nowacka, M. (2010). The ultimate attainment of English pronunciation by Polish college students: A longitudinal study. In E. Waniek-Klimczak (ed.), *Issues in accents of English 2: Variability and Norm.* Newcastle, UK: Cambridge Scholars Publishing.

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*, 175-184.

Rohena-Madrazo, M. (2013). Perceptual assimilation of occluded voiced stops by Spanish listeners. *Selected Proceedings of the 15th Hispanic Linguistics Symposium 2013* (pp. 140-156).

Rojczyk, A. (2009). Modele percepcji systemu dźwiękowego języka obcego. In J. Nijakowska (ed.), *Język, poznanie, zachowanie: Perspektywy i wyzwania w studiach nad przyswajaniem języka obcego* (pp. 120-135). Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Rojczyk, A. (2010). Forming new vowel categories in second language speech: The case of Polish learners' production of English /ɪ/ and /e/. *Research in Language, 8,* 85-97.

Rojczyk, A. (2011). Overreliance on duration in nonnative vowel production and perception: The within lax vowel category contrast. In M. Wrembel, M. Kul & K. Dziubalska-Kołaczyk (eds.), *Achievements and Perspectives in SLA of Speech: New Sounds 2010, Volume II* (pp. 239-250). Frankfurt am Main: Peter Lang.

Sobkowiak, W. (2004). *English phonetics for Poles.* Poznań, Poland: Wydawnictwo Poznańskie.

Wong, J. W. S. (2015). Perceptual training intensity effects on the perception and production of English /ɪ/ and /iː/ by Cantonese ESL learners. In *Proceedings of 4th International Conference on English Pronunciation: Issues and Practices* (pdf version) (pp. 161-163).

To cite this article: Lipińska, D. (2017). Production of English /iː/ and /ɪ/ by Polish learners of English: A longitudinal study. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 180-185). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Vowel reduction in the interlanguage phonology of Bulgarian learners of Modern Greek – An acoustic study

Milena Milenova
milenova.m@gmail.com

Sofia University St. Kliment Ochridksi, Aristotle University of Thessaloniki

Abstract. This study is part of a larger project investigating the production of unstressed /o/ and /a/ by Bulgarian learners of Modern Greek. This paper reports on the results for post-stressed /o/. The /o/ vowel exists in both the native language (NL) and the target language (TL), but it is subject to a different pattern of vowel reduction in each. In Contemporary Standard Bulgarian (CSB) unstressed /o/ undergoes phonological vowel reduction: its vowel quality is neutralised as it is merged with /u/. In Standard Modern Greek (SMG), the reduction is phonetic: unstressed /o/ is raised but its distinct vowel quality is retained. This study explores the transfer of NL phonological vowel reduction. Its objective is to investigate the effect of pronunciation training on the acquisition of the TL phonetic vowel reduction by beginner learners. To this end, an acoustic experiment was conducted. The experimental group of the study was composed of 7 native female speakers of CSB. A group of 3 native female speakers of SMG served as a control. The learners were recorded twice: prior to pronunciation instruction and after 50 one-hour, pronunciation-instruction sessions embedded in core Modern Greek lessons. The elicitation protocol of the experiment was composed of the symmetrical disyllable [ˈpVpV] with V=/o, u/ embedded in a carrier sentence. The data were hand-annotated and analysed using Praat (Boersma & Weenink 2017). The frequencies of F1 and F2 were measured at the vowel mid-point. *T*-tests were performed to compare the process of vowel reduction in CSB and SMG and to examine the realisation pattern of post-stressed /o/ employed by the learners prior to and after teaching intervention. The analyses of the data revealed that the target vowel reduction pattern which retains the vowel quality of post-stressed /o/ was not acquired by the learners. The findings of the study suggest that the acquisition of the TL vowel reduction pattern may be unfeasible for the novice learner.

Keywords: NL transfer, vowel reduction, Modern Greek, Bulgarian

Introduction

The acquisition of a non-native language vowel reduction pattern is reportedly a challenging task for second language learners. The transfer of the native language vowel reduction pattern seems a common phenomenon in interlanguage phonologies. Russian learners of English systematically apply the rules of excessive Russian vowel reduction in their speech in the target language (Banzina, 2012). Spanish speakers whose L1 exhibits vowel reduction as formant undershoot are reported to produce incorrectly the reduced unstressed English vowels (Flege & Bohn, 1989; Gutiérrez & Monroy, 2003; Gutiérrez, 2004; Hammond, 1986). Similarly, Dutch speakers transfer their native vowel reduction pattern to the production of English words containing reduced unstressed vowels (Braun, Lemhöfer, & Cutler, 2008). Bulgarian Judeo-Spanish displays the vowel reduction rules of raising unstressed /a/ and /o/ to /ɜ/ and /u/, respectively. Notably this pattern of vowel reduction is found only in the variety of Judeo-Spanish spoken by Bulgarian Spanish bilinguals due to transfer from their dominant language (Gabriel & Kireva, 2014).

The present author's impressionistic observations on the oral productions of Bulgarian speakers of Modern Greek suggest that the extreme reduction of unstressed /o/ and its realisation as /u/ is a persistent feature of their interlanguage phonology. Presumably this mispronunciation is caused by the transfer of the NL vowel reduction pattern to the TL.

Background

Vowel reduction involves shortening of the duration of vowels and changes in their formant structure. The process is conditioned by factors like stress, speech tempo and context. Reduced vowels occupy different positions in the F1xF2 acoustic space with respect to their stressed counterparts. The degree of reduction varies cross-linguistically. Depending on the language, the displacement of unstressed vowels may involve centralization in the F1xF2 space or neutralisation of vowel qualities. The former pattern of vowel reduction is referred to as phonetic vowel reduction, whereas the latter as phonological vowel reduction (Crosswhite, 2001; Fourakis, 1991). Phonological vowel reduction refers to categorical substitution of sounds. Phonetic vowel reduction results from undershoot of vowel targets (Padgett & Tabain, 2005). Contemporary Standard Bulgarian (CSB) and Standard Modern Greek (SMG) exhibit phonological and phonetic vowel reduction respectively.

Vowel reduction in CSB

Vowel reduction is one of the most salient features of the vowel system in CSB. In stressed syllables, CSB contrasts 6 vowels /i, ɛ, a, ɜ, ɔ, u/. There is a strong differentiation between stressed and unstressed syllables (Boyadzhiev & Tilkov, 1999). In unstressed positions all vowels except /i/ are raised, /a/ and /ɔ/ are neutralised to /ɜ/ and /u/, and the 6-way contrast is reduced to the 4-way one: /i, ɛ, ɜ, u/ (Andreeva, Barry, & Koreman, 2014). The process is obligatory and does not depend on speech tempo (Boyadzhiev & Tilkov, 1999; Crosswhite, 2001; Zhobov, 2004). The changes in vowel quality are position sensitive: post-stressed vowels undergo more profound quality changes than pre-stressed ones (Boyadzhiev & Tilkov, 1999; Zhobov, 2004). It has been suggested that the reduction of unstressed vowels is the result of reduced mandibular aperture (Wood & Pettersson, 1988).

Vowel reduction in SMG

The vowel system of SMG comprises 5 vowel categories: /i, e, a, o, u/. Unstressed vowels tend to centralise and occupy a reduced vowel space compared to their stressed counterparts. Studies show that the process applies under both spontaneous and controlled speech tempo conditions (Baltazani, 2007; Fourakis, Botinis & Katsaiti, 1999; Nicolaidis, 2003). Unlike vowel reduction in CSB, the process is not mandatory. In SMG vowel reduction is phonetic, i.e. the lack of stress results in formant undershoot, but it does not lead to merging of vowel categories (Arvaniti, 2007).

Aims and method

Aims

The current study explores the acquisition of phonetic vowel reduction by learners whose NL applies phonological vowel reduction. The aim is to investigate whether novice Bulgarian learners of SMG who study the language in formal settings will acquire the TL articulation pattern which retains the distinctive category of /o/ in unaccented syllables. Specifically, the primary objective is to explore whether explicit pronunciation training is beneficial for the acquisition of the TL vowel reduction pattern. To the author’s knowledge, no previous acoustic study has compared either the vowel system of CSB and SMG, or the process of vowel reduction in these two languages. Thus, the secondary objective of this study is to examine the acoustic properties, namely F1 and F2, of the vowel categories /o/ and /u/ in CSB and SMG, and to explore the reduction of /o/ in post-stressed syllables in the two languages.

For the purpose of the study a two-session acoustic experiment was conducted to investigate the formant structure of the target vowel produced by the learners before and after training, and to compare it to the productions of the native speakers of SMG.

Participants

Seven native female speakers of CSB (M_{age}=18.85, s.d.=0.37), beginner learners of SMG enrolled in a Modern Greek Studies programme at the University of Sofia, comprised the experimental group of the study. A group of 3 native female speakers of SMG (3 females M_{age}= 28, s.d.=2.3) served as a control. The recordings of the experimental group were made in a sound treated interpreter booth at the University of Sofia. The recordings of the control group were made in a quiet room in an apartment in Athens. The exact purpose of the experiment was not made clear to the participants. The learners were recorded twice: prior to the pronunciation instruction (T1) and after the pronunciation instruction sessions (T2). At T1, the learners had received approximately 200 hours of formal instruction without any training on the articulation of unstressed /o/ in SMG. At T2, the students had received another 300 hours of formal instruction which included pronunciation training.

Pronunciation training

The pronunciation training sessions were embedded in the core SMG lessons. The students received a total of 50 teaching hours of instruction dealing with different aspects of the pronunciation of SMG. The pronunciation instruction for unstressed /o/ included awareness raising activities and production training. The former focused on providing the students with the information about the different patterns of realisation of unstressed /o/ in SMG and CSB. The production instruction focused on training the learners to produce unstressed /o/ with larger mandibular opening. The target articulation was practised in a variety of controlled and free oral production activities. The students were provided with individual feedback.

Stimuli and procedure

The elicitation protocol of the experiment was composed of the disyllables [ˈpopo] and [ˈpupu]. The particular bilabial phonetic context was chosen since the articulation of /p/ does not involve movement of the tongue thus excluding interference with lingual gestures and co-articulation effects on vowel quality (Sfakianaki, 2012:193). The stimuli were embedded in the carrier sentence: ‘ˈleʝe _______ ˈpali’ (‘Say ________ again’). The speech material was recorded digitally at 44.1 kHz sampling rate using a portable Marantz PMD 660 Solid State recorder and a RØDE NTSS condenser microphone placed at approximately 45-degree angle and 20 cm away from the corner of the speaker’s mouth to avoid overload. Ten repetitions of the carrier phrase were recorded and eight of them were used for the analysis.

Analyses

The data were hand annotated and analysed using Praat (Boersma & Weenink, 2017). The corpus of the experimental group comprised 448 tokens (7 speakers x 8 repetitions x 2 vowels x 2 stress conditions x 2 recordings) rendering 896 formant measurements. The corpus of the control group comprised 96 tokens (3 speakers x 8 repetitions x 2 vowels x 2 stress conditions) rendering 192 formant measurements. The frequencies of F1 and F2 were measured at the vowel mid-point using a 25 ms Gaussian window and preemphasis of 50 Hz. The measurements were obtained using a script running in Praat. Whenever needed, the formant values were manually corrected following the correcting procedure described in Sfakianaki (2012:155-157). Independent and paired *t*-tests were performed to calculate the statistical significance of the results.

Results

Table 1 displays the mean F1 and F2 values of the vowels produced by the control group (CG) and the experimental group prior to the pronunciation training (EG T1) and after the teaching intervention (EG T2).

Table 1. Mean values of F1 and F2 of stressed and post-stressed /o/ and /u/ split by group and time of production

Vowel	F1, Hz			F2, Hz		
	EG T1	EG T2	CG	EG T1	EG T2	CG
/o/$_{[+str]}$	565	543	529	984	989	918
/o/$_{[-str]}$	422	403	473	831	856	841
/u/$_{[+str]}$	410	413	378	761	781	765
/u/$_{[-str]}$	400	409	384	844	863	724

The values displayed in Table 1 suggest that in terms of formant structure CSB /o/ and SMG /o/ are similar vowels. The same refers to CSB /u/ and SMG /u/. Independent samples *t*-tests were performed to estimate the degree of similarity of these vowels with respect to F1 and F2 in stressed and post-stressed syllables. The significance results of the pairwise comparisons are reported in Table 2.

Table 2. Significance results of the pairwise comparisons of F1 and F2 of CSB & SMG /o/ and /u/

Formant	*t*	*df*	*p*-value	*r*
F1 bg/o/$_{[+str]}$ vs. gr/o/$_{[+str]}$	2.4	78	.019	0.26
F2 bg/o/$_{[+str]}$ vs. gr/o/$_{[+str]}$	3.6	78	.001	0.38
F1 bg/o/$_{[-str]}$ vs. gr/o/$_{[-str]}$	-3.08	29	.004	0.49
F2 bg/o/$_{[-str]}$ vs. gr/o/$_{[-str]}$	-.441	78	.661	0.05
F1 bg/u/$_{[+str]}$ vs. gr/u/$_{[+str]}$	3.09	78	.003	0.33
F2 bg/u/$_{[+str]}$ vs. gr/u/$_{[+str]}$	-.2	78	.837	0.02
F1 bg/u/$_{[-str]}$ vs. gr/u/$_{[-str]}$	1.4	31	.181	0.23
F2 bg/u/$_{[-str]}$ vs. gr/u/$_{[-str]}$	4.5	77	.000	0.45

Table 3. Significance results of the pairwise comparisons of /o/ and /u/ produced by the CG

Formant	*t*	*df*	*p*-value	*r*
F1 gr/o/$_{[+str]}$ vs. gr/o/$_{[-str]}$	4	23	.001	0.63
F2 gr/o/$_{[+str]}$ vs. gr/o/$_{[-str]}$	6	23	.000	0.75
F1 gr/o/$_{[-str]}$ vs. gr/u/$_{[+str]}$	8	23	.000	0.84
F2 gr/o/$_{[-str]}$ vs. gr/u/$_{[+str]}$	5	23	.000	0.69
F1 gr/o/$_{[-str]}$ vs. gr/u/$_{[-str]}$	9	23	.000	0.87
F2 gr/o/$_{[-str]}$ vs. gr/u/$_{[-str]}$	7	23	.000	0.81
F1 gr/u/$_{[+str]}$ vs. gr/u/$_{[-str]}$	-.6	23	.555	0.12
F2 gr/u/$_{[+str]}$ vs. gr/u/$_{[-str]}$	2.6	23	.018	0.47

Considering the significance values and their respective effect sizes the results reveal that SMG /o/ is slightly higher and moderately retracted compared to its CSB counterpart. In post-stressed position the two vowels are distinct only in terms of height, suggesting that the raising of CSB post-tonic /o/ is more profound, as can be observed by the effect size of this difference. This indicates that SMG post-stressed /o/ is produced with more open articulation compared to its CSB counterpart. As for the /u/ vowel the results reveal that SMG /u/ and CSB /u/ differ only in terms of height, the latter being moderately lower. Their unstressed counterparts differ only in terms of backness: SMG unstressed /u/ is considerably more retracted than CSB unstressed /u/.

To further explore the target vowel reduction pattern paired *t*-tests were conducted to estimate the degree of change in the formant structure of post-stressed /o/ in SMG. The results are demonstrated in Table 3.

Table 4. Significance results of the pairwise comparisons of the vowels /o/ and /u/ produced by the EG at T1 and T2

Formant	*t*	*df*	*p*-value	*r*
T1 F1: bg/o/$_{[+str]}$ vs. bg/o/$_{[-str]}$	16	55	.000	0.90
T2 F1 bg/o/$_{[+str]}$ vs. bg/o/$_{[-str]}$	14	55	.000	0.88
T1 F2 bg/o/$_{[+str]}$ vs. bg/o/$_{[-str]}$	11	55	.000	0.83
T2 F2 bg/o/$_{[+str]}$ vs. bg/o/$_{[-str]}$	6	55	.000	0.63
T1 F1 bg/o/$_{[-str]}$ vs. bg/u/$_{[+str]}$	1.5	55	.131	0.20
T2 F1 bg/o/$_{[-str]}$ vs. bg/u/$_{[+str]}$	-1.5	55	.148	0.19
T1 F2 bg/o/$_{[-str]}$ vs. bg/u/$_{[+str]}$	4.4	55	.000	0.51
T2 F2 bg/o/$_{[-str]}$ vs. bg/u/$_{[+str]}$	2.4	55	.022	0.30
T1 F1 bg/o/$_{[-str]}$ vs. bg/u/$_{[-str]}$	2.4	55	.018	0.31
T2 F1 bg/o/$_{[-str]}$ vs. bg/u/$_{[-str]}$	-.7	55	.517	0.09
T1 F2 bg/o/$_{[-str]}$ vs. bg/u/$_{[-str]}$	.48	55	.633	0.06
T2 F2 bg/o/$_{[-str]}$ vs. bg/u/$_{[-str]}$	-.4	55	.72	0.05
T1 F1 bg/u/$_{[+str]}$ vs. bg/u/$_{[-str]}$	.81	55	.421	0.10
T2 F1 bg/u/$_{[+str]}$ vs. bg/u/$_{[-str]}$	.51	55	.61	0.07
T1 F2 bg/u/$_{[+str]}$ vs. bg/u/$_{[-str]}$	26	55	.000	0.99
T2 F2 bg/u/$_{[+str]}$ vs. bg/u/$_{[-str]}$	-2.8	55	.007	0.35

In line with previous research (see Arvaniti, 2007 for a review) the results reveal that in SMG unstressed /o/ is significantly raised but remains distinct from both stressed and unstressed /u/. To examine the effect of pronunciation training on the acquisition of the TL production of post-stressed /o/ by the experimental group paired samples *t*-test were performed. The significance of the pairwise comparisons is reported in Table 4. The results reveal that at T1 post-stressed /o/ is not distinct from stressed /u/ in terms of F1 and not distinct from post-stressed /u/ in terms of F2. It differs only moderately from post-stressed /u/ in terms of F1. At T2 post-stressed /o/ produced by the learners does not differ from post-stressed /u/.

These findings indicate that the target realisation pattern of unstressed /o/ was not acquired after the teaching intervention. At T2 the learners continued to apply a less open articulation merging post-stressed /o/ with post-stressed /u/. All the results discussed so far are exemplified in the F1xF2 plots in Figure 1.

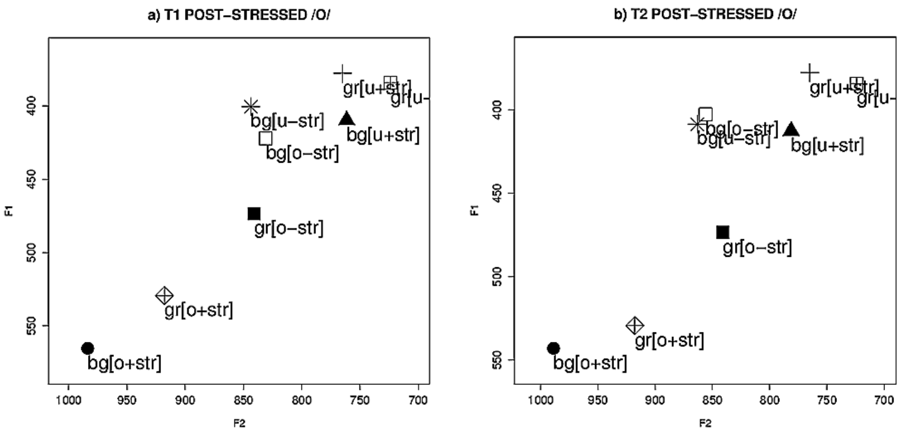


Figure 1a, b. Plots of /o/ and /u/ produced in the disyllable [ˈpVpV] by the CG and the EG at T1 and T2

The different reduction patterns of post-stressed /o/ employed by the EG and the CG are clearly demonstrated in the vowel plots. In both CSB and SMG, post-stressed /o/ is raised but the degree of raising is considerably different in the two languages. It can be seen that the pattern of realisation of the post-tonic /o/ by the EG is the same at T1 and T2. At both the pre-test and the post-test, the unstressed /o/ produced by the EG is less open than the unstressed /o/ produced by the native speakers suggesting that the learners' articulation was not improved after teaching intervention.

Discussion and conclusions

This experimental study set out to explore the effect of pronunciation training on the acquisition of phonetic vowel reduction by beginner Bulgarian learners of Modern Greek whose NL displays phonological vowel reduction. The results revealed that after training the learners continued to produce post-stressed /o/ employing the NL vowel reduction pattern which neutralises its vowel quality. A tentative interpretation of this finding could be that the suppression of the transfer of the NL phonological vowel reduction may be too demanding a task for the novice learner. Lack of training effect on the acquisition of English phonological vowel reduction by Spanish learners, whose NL displays phonetic vowel reduction, has been reported by Gutiérrez and Monroy (2003) and Gutiérrez (2004). The authors concluded that the brevity of the training period (7 hours) was the main reason for the absence of pronunciation improvement in the investigated groups of Spanish speakers. The learners in the present study received a total of 50 hours of pronunciation instruction distributed over the time-span of 7 months. This implies that the lack of training effect could be due to other factors such as the particular training procedure.

Gutiérrez (2004) suggested that the lack of effect of treatment observed in his study could be due to the inclusion of other aspects of English pronunciation in the training sessions. Similarly, in our case the focus of the pronunciation lessons was not exclusively on the production of unstressed /o/, but also included other segmental and suprasegmental aspects of the TL. This could have rendered the training procedure inefficient to block the transfer of the NL vowel reduction pattern. The current study has several limitations which provide implications for future research. First, the small number of participants precludes generalizability of the findings. Second, the acquisition of the TL vowel reduction pattern was assessed by investigating only a group of beginner learners. Further investigation should be carried out with intermediate, advanced and proficient groups of learners in order to explore the effect of TL competence level on the suppression of the NL vowel reduction transfer. In addition, the assessment of the effect of different types of pronunciation training could shed more light on their efficacy for the acquisition of the TL vowel reduction.

References

Andreeva, B., Barry, W., & Koreman, J. (2013). The Bulgarian stressed and unstressed vowel system. A corpus study. *14th annual conference of the international speech communication association, Interspeech 2013*, August 26-29, Lyon, France, 345-348.

Arvaniti, A. (2007). Greek phonetics: the state of the art. *Journal of Greek Linguistics, 8*, 97-208.

Baltazani, M. (2007). Prosodic rhythm and the status of vowel reduction in Greek. *Selected Papers on Theoretical and Applied Linguistics from the 17th International Symposium on Theoretical & Applied Linguistics* (Vol. 1, 31-43). Department of Theoretical and Applied Linguistics, Thessaloniki.

Banzina, E. (2012). *The role of secondary-stressed and unstressed-unreduced syllables in word recognition: Acoustic and perceptual studies with Russian learners of English*. PhD Thesis. Bowling Green State University.

Boersma, P., & Weenink, D. (2017). *Praat: Doing phonetics by computer* (Version 6.0.29) [Computer program]. Retrieved June 15, 2017, from http://www.praat.org.

Boyadzhiev, T., & D. Tilkov. 1999. *Phonetica na bylgarskiya ezik*. Veliko Tyrnovo: Abagar.

Braun, B., Lemhöfer, K., & Cutler, A. (2008). English word stress as produced by English and Dutch speakers: The role of segmental and suprasegmental differences. In *Proceedings of Interspeech 2008* (pp. 1953-1953).

Crosswhite, K. (2001). *Vowel reduction in optimality theory*. New York, NY: Routledge.

Flege, J. E., & Bohn, O.S. (1989). An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition*, *11*, 35-62.

Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *Journal of the Acoustical Society of America*, *90*, 1816-1827.

Fourakis, M. A., Botinis, & M. Katsaiti. (1999). Acoustic characteristics of Greek vowels. *Phonetica*, *56*, 28-43.

Gabriel, C., & Kireva, E. (2014). Speech rhythm and vowel raising in Bulgarian Judeo-Spanish. In N. Campbell, D. Gibbon & D. Hirst (eds.), *Proceedings of the 7th international conference on speech prosody 7* (pp. 728-732).

Gutiérrez, F. (2004). Acento y debilitamiento vocálico en el inglés de un grupo de estudiantes españoles de Enseñanza Secundaria Obliagoria (ESO). *Porta Linguarum*, *2*, 111-124.

Gutiérrez, F., & Monroy, R. (2003). Contraste entre las vocales átonas del inglés y el español: El debilitamiento vocálico en el inglés de un grupo de hispanohablantes. *Proceedings of the 3rd International Contrastive Linguistics Conference* (pp. 211-222). September 2003.

Hammond, R. H. (1986). Error analysis and the natural approach to teaching foreign languages. *Lenguas Modernas*, *13*, 129-139.

Nicolaidis, K. (2003). Acoustic variability of vowels in Greek spontaneous speech. *Proceedings of the XVth international congress of phonetic science*s (pp. 3221-3224). Autonomous University of Barcelona.

Sfakianaki, A. (2012). *An acoustic study of coarticulation in the speech of Greek adults with normal hearing and hearing impairment*. PhD Thesis. Aristotle University of Thessaloniki.

Padgett, J., & Tabain, M. (2005). Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica*, *62*, 14-54.

Wood, S., & Pettersson, T. (1988). Vowel reduction in Bulgarian: the phonetic data and model experiments. *Folia linguistica*, Tomus XXII, 239-262.

Zhobov, V. (2004). *Zvukovete v bylgarskiya ezik*. Sofia: Sema.

To cite this article: Milenova, M.. (2017). Vowel reduction in the interlanguage phonology of Bulgarian learners of Modern Greek - An acoustic study. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 186-192). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

On the mixed nature of L3 Spanish grammar of L1 Japanese subjects with L2 English

Yoichi Miyamoto[1], Kazumi Yamada[2]
miyamoto@lang.osaka-u.ac.jp, kzyamada@kwansei.ac.jp

[1]Osaka University, [2]Kwansei Gakuin University

Abstract. We examine how first language (L1) Japanese subjects with L2 English acquire null arguments in their L3 Spanish. The results from an acceptability judgment task and a truth value judgment task show: i) that intermediate subjects allow both strict and sloppy readings with null subjects and objects and ii) lower and upper advanced subjects permit only a strict reading with null subjects, but still allow both strict and sloppy readings with null objects. We argue that the finding in (i) follows from L1 transfer, bypassing their L2 English, which prohibits null arguments altogether, and that the finding in (ii) results from L3 acquisition being morphosyntactically based. Oku (1998) demonstrates convincingly that Japanese null arguments permit both sloppy and strict readings. Given that Spanish null subjects without a clitic only permit strict reading (Duguine, 2013; Oku, 1998, among others), this dissipation of the sloppy reading with null subjects at the lower advanced level suggests that they successfully acquire Spanish null subjects. Conversely, although Spanish null objects without the object clitic *lo* refer to not only [-definite, -specific] but also [+definite, +specific] objects, if the referent is identifiable from the immediate discourse (Campos, 1986; Masullo, 2003, among others), statistical analyses reveal that lower and upper advanced subjects act differently. Null objects at the lower advanced and intermediate levels are more like their Japanese counterparts, while ones at the upper advanced level are more like the ones Spanish native speakers use. Significantly, only the upper advanced subjects have also mastered object clitics, which causes us to speculate that object clitics may act as a trigger to acquire Spanish null objects.

Keywords: L3 Spanish, L2 English, L1 Japanese, null subject, null object, object clitic

Introduction

The current paper reports experimental data showing that L1 Japanese/L2 English learners of L3 Spanish can acquire Spanish null arguments (NAs) at an advanced level. We also show that the learners acquire subject NAs earlier than object NAs. Of additional significance is the finding that only those learners who acquire object clitics acquire object NAs, which suggests that acquisition of object NAs hinges on acquisition of object clitics. In light of our results, we evaluate four representative L3 acquisition models.

Theoretical background on null arguments

Japanese null arguments

Recent developments in Japanese syntax reveal that Japanese NAs are not pronominal in nature. Strong supportive evidence stems from the fact that Japanese NAs permit both strict and sloppy readings, as Oku (1998) shows. Accordingly, (1b), which follows (1a), is ambiguous.

(1) a. Mary-wa [zibun-no ronbun-ga saiyoosareru]-to omotteiru.
Mary-TOP self -GEN paper -NOM will be accepted-that think
'Mary thinks that her paper will be accepted.'
b. John-mo [[e] saiyoosareru]-to omotteiru.
John-also will be accepted-that think
'John also thinks that [e] will be accepted.' (Oku, 1998:166)

(2) a. Mary thinks that her paper will be accepted. b. John also thinks that it will be accepted.
In (1b), the null embedded subject [e] can be interpreted not only as Mary's paper (strict reading), but also John's (sloppy reading). If Japanese NAs are covert pronominals, the embedded subject in (1b) should allow for only a strict reading as Mary's paper. In (2b) the pronoun *it* only allows the strict reading, which is as expected. The contrast between (1b) and (2b) therefore shows that Japanese NAs are not null pronouns (see also Saito, 2007).

Spanish null arguments

In contrast to their Japanese counterparts, Spanish NAs appear mixed in nature. When it comes to subject NAs only the strict reading is available (see Duguine, 2013; Oku, 1998 for discussion).

(3) a. Maria cree [que sue propuesta sera aceptada] y
Maria believes that her proposal will-be accepted and
'Mary thinks that her paper will be accepted.'
b. Juan tambien cree [[e] sera aceptada].
Juan too believes will-be accepted
'Juan also thinks that [e] will be accepted.' (Oku, 1998: 166)

The embedded subject NA [e] in (3b) can only be interpreted as Maria's paper.

In contrast, Campos (1986) claims that object NAs without the clitic *lo* only permit [-definite, -specific] reading. However, there are some intervening factors for his claim. Masullo (2003) shows that if the referent is recoverable from the immediate context, the object NA without the object clitic does not have to be [-definite, -specific]. The example in (4), for example, is acceptable even without an object clitic:

(4) two persons leaving a room, one says to the other.
Apaga [e] [i.e. la luz, la televisión, etc.]
turn off the light the TV (Alamillo & Schwenter, 2007:113)

In short, subject NAs without the clitic only allow a strict reading, whereas object NAs without the clitic are permitted in the [-definite, -specific] context, as well as in the [+definite, +specific] context, if the referent is identifiable from the context. If this is correct, it would be unsurprising if the object NAs permit a sloppy reading, in addition to a strict reading, if the context is appropriate.

L3 acquisition models

Four models of L3 acquisition have been proposed in the literature. Notice that in the case of L3, we have two languages, L1 and L2, to potentially make impact on the initial stages of L3 acquisition, and make predictions on L3 development.

The first model is the absolute L1 transfer model (e.g., Na Ranong & Leung, 2009), which suggests that L1 necessarily influences L3 development. The second model is the L2 status factor model (e.g., Bardel & Falk, 2007). According to this second model, it is the L2 which impacts on L3 development. The third and fourth models share the idea that both L1 and L2 potentially influence L3 development; but, the second, third and fourth, diverge in the way they do. The third model, named the cumulative enhancement model (Flynn, Foley, & Vinnitskaya, 2004), states that the L1 or the L2 which facilitates L3 development influences L3 development. In the case that neither the L1 nor the L2 facilitates L3 development, they would be silent on L3 development. The last model, which is dubbed the typological primacy model (Rothman, 2010), suggests that L1 or L2, typologically closer to L3, influences L3 development.

In light of these four L3 acquisition models, we turn to the current experiment on the acquisition of Spanish NAs by L1 Japanese/L2 English subjects.

Experiment

Predictions

Recall that Japanese NAs allow both strict and sloppy readings, whereas in Spanish, subject NAs only permit a strict reading and object NAs may allow both readings if the context is appropriate. On this basis, the four L3 acquisition models allow for the following predictions: the L1 transfer model - NAs in their L3 Spanish should first permit both strict and sloppy readings; the L2 status factor model - No NAs in their L3 Spanish should be permitted if their L2 is at advanced level, given Miyamoto and Yamada's (2015) finding that L1 Japanese speakers stop accepting NAs in their L2 English at the advanced level; the cumulative enhancement model - NAs in their L3 Spanish may first allow both readings due to L1 transfer since NAs are available in both Japanese and Spanish, but not in English; and, the typological primacy model - since typologically speaking, Spanish and English are grouped together, no NAs may be available in their L3 Spanish due to L2 influence, again, if their L2 English is at advanced level.

Participants

A total of 33 subjects participated in our study. The experimental group consisted of 23 L1 Japanese/L2 English learners of L3 Spanish, further divided into two proficiency groups: 13 advanced and 10 intermediate learners. The advanced learners were 3rd and 4th year students in a Japanese university, majoring in Spanish, or were CEFR C1 or B2 successful candidates. The participants' L2 English proficiency was considered advanced, based on the fact that they passed the university English entrance examination, consisting mostly of written responses of between 80 to 100 words, or more. Although we do not have access to participants' precise English scores, given that their English scores constitute 50% of the total scores, a poor English test score would have resulted in a failure to pass the university entrance examination. The intermediate learners of L3 Spanish, on the other hand, were 2nd or 3rd year students in a different Japanese university, taking Spanish language class as an elective subject. Their TOEIC scores ranged from 525 to 670, or they passed the level 2 or pre-level 2 English Language Proficiency test, their English proficiency level was intermediate. Ten Spanish native speakers joined our study as a control group.

Stimuli and procedures

The L3 learners were tested with a truth-value judgment task (TVJT) and an acceptability judgment task (AJT). First, the TVJT was conducted in order to investigate the availability of sloppy and strict readings with NAs. Each stimulus consisted of a dialogue in Japanese, followed by a Spanish test sentence. The learners were then required to judge whether test sentences correctly described the situations of given dialogues by circling 'correct' or 'incorrect'. The TVJT consisted of 52 stimuli, including 28 sentence types where four sentence types matched the current study: subject NA sloppy, subject NA strict, object NA sloppy, and object NA strict, each has two tokens. A sample test item is given in (5). The dialogue is translated into English for convenience.

(5) Sentence type: object NA sloppy

1.

My car is very dirty.

2.

It's very clean now.

3.

I should clean the car, too

4.

Now, it is very clean.

Test sentence: Oso lavó su propio coche. Y Pingüino también limpió [e].
'Bear cleaned his own car and Penguin cleaned [e], as well.' Correct / Incorrect

The second task was the AJT in which L3 learners were asked to judge whether each Spanish sentence including a subject NA or an object NA is correct or not. The task consisted of six stimuli:

three included subject NAs and three object NAs. A sample object NA item is shown in (6).

(6) Tarou estropeó la computadora pero su padre la arregló.
'Although Taro broke a computer, his father fixed [it].'

The subjects were also asked to correct sentences they judged unacceptable. Note that no clitic was involved in the object NA test sentences without any context so that no object NAs should be allowed.

Results

We first report the results of the AJT to explore the extent to which the L3 learners allowed NAs in their L3 grammar. All L3 learners allowed subject NAs. Moreover, all intermediate L3 learners permitted object NAs. By contrast, among the 13 advanced learners, nine learners allowed object NAs while four learners did not. Significantly, they noted that the test sentences with an object NA should be accompanied by the object clitic, indicating that they must have acquired the object clitic. We therefore divide the advanced learners into the lower and upper advanced learners.

As Table 1 shows, no preference was observed between the three items in each condition.

Table 1. Descriptive item analysis (J-SFL learners, n=23)

Condition	Level		Acceptance	N	Item 1	Item 2	Item 3
Subj NAs	Upper Advanced	(n=4)	91.7%	11/12	4	4	3
	Lower Advanced	(n=9)	85.2%	23/27	9	6	8
	Intermediate	(n=10)	96.7%	29/30	10	9	10
Obj NAs	Upper Advanced	(n=4)	0%	0/12	0	0	0
	Lower Advanced	(n=9)	66.7%	18/27	2	8	8
	Intermediate	(n=10)	96.7%	29/30	10	9	10

The TVJT results are summarized in Table 2, and the statistical analysis in Table 3.

Table 2. Acceptance rates - subject NA and object NA (%)

		Subject NA		Object NA	
		Sloppy	Strict	Sloppy	Strict
Control	(n=10)	5.0	70.0	50.0	90.0
Upper Advanced	(n=4)	0	87.5	75.0	75.0
Lower Advanced	(n=9)	22.2	94.4	94.4	38.9
Intermediate	(n=10)	90.0	70.0	75.0	70.0

Table 3: One-way between group ANOVA (independent)

Condition	Main Effect	Between Groups (compared with Control)		
Subj NA sloppy	$F(3,29) = 26.24$, $p<0.001$	U-Adv ($p>0.05$)	L-Adv ($p>0.05$)	Int ($p<0.001$)
Subj NA strict	No main effect	No significance		
Obj NA sloppy	$F(3,29) = 2.96$, $p<0.05$	U-Adv ($p>0.05$)	L-Adv ($p<0.05$)	Int ($p>0.05$)
Obj NA strict	$F(3,29) = 3.92$, $p<0.05$	U-Adv ($p>0.05$)	L-Adv ($P<0.05$)	Int ($p>0.05$)

U-Adv = Upper Advanced; L-Adv = Lower Advanced; Int = Intermediate

Finally, Table 4 shows an item analysis, conducted because of the relatively small number of items (i.e., two items) per condition in TVJT. Among the J-SFL's results, no preference was observed between the two items in each condition, as the shaded columns indicate.

Table 4: Descriptive item analysis (J-SFL learners, n=23)

Condition	Level		Acceptance	N	Item 1	Item 2
Subj NA sloppy	Upper Advanced	(n=4)	0%	0/8	0	0
	Lower Advanced	(n=9)	22.2%	4/18	2	2
	Intermediate	(n=10)	90.0%	18/20	8	10
Subj NA strict	Upper Advanced	(n=4)	87.5%	7/8	3	4
	Lower Advanced	(n=9)	94.4%	17/18	8	9
	Intermediate	(n=10)	70.0%	14/20	7	7
Obj NA sloppy	Upper Advanced	(n=4)	75.0’	6/8	3	3
	Lower Advanced	(n=9)	94.4%	17/18	9	8
	Intermediate	(n=10)	75.0%	15/20	10	5
Obj NA strict	Upper Advanced	(n=4)	75.0%	6/8	3	3
	Lower Advanced	(n=9)	38.9%	7/18	2	5
	Intermediate	(n=10)	70.0%	14/20	8	6

To summarize, in the subject NA sloppy context, there was no significant difference in performance between the control group and both advanced groups. In the object NA context, there was no significant difference in performance between the control and the upper advanced group, however, interestingly, a significant difference was observed between the control group and the lower advanced group. This indicates that the L3 grammar of the upper advanced group is close to Spanish native speakers’ grammar.

Discussion

This section is devoted to a brief discussion outlining our main findings.

First, the fact that the intermediate subjects permitted both strict and sloppy readings is consistent with the all four L3 acquisition models, because their L2 English has not reached the advanced level (see Miyamoto & Yamada, 2015 for the claim that it is at the advanced level that no NAs are accepted in the L2 English grammar of L1 Japanese speakers), and thus, it is very likely that their L2 is influenced by L1 Japanese. In other words, irrespective of whether the L1 or L2 is transferred, the L3 Spanish is influenced by the L1.

Conversely, it is worth examining the behavior of the advanced L3 learners, who also have advanced English proficiency. First, to reiterate, irrespective of whether object clitics are acquired, the advanced subjects only permitted the strict reading with subject NAs in the L3 Spanish. This indicates that Spanish subject NAs are acquired at the lower advanced level. By contrast, it is significant that only those learners who have mastered object clitics have a good command of object NAs. While object NAs can be ultimately acquired, there is a time lag between subject and object NAs.

Here we need to focus on the results on object NAs from the early advanced subjects. Since they have a good command of English, if the L2 influence had occurred, they should not have accepted any NAs to begin with. The fact that they accepted NAs with strict and sloppy readings, therefore, indicates that their L3 is influenced by the L1 Japanese. Consequently, the current study supports the L1 transfer model and the cumulative enhancement model, both of which predict the L1 transfer. The obvious difference between the upper and lower advanced subjects is whether object clitics are mastered, and therefore, we suggest that advanced subjects take object clitics to be a trigger for the acquisition of object NAs. Since it has been shown that object clitics are difficult to acquire (e.g., Castilla & Pérez-Leroux, 2010), it would be unsurprising that it takes more time to acquire object NAs than subject NAs.

Conclusion

The current paper reports that the L3 Spanish grammar of L1 Japanese/L2 English learners is influenced by their L1, and supports the absolute L1 transfer model and the cumulative enhancement model in the sense that these two models predict L1 transfer. We also show that object NAs are acquired after subject NAs. We further highlight that the acquisition of object NAs is connected to the acquisition of object clitics; only those who have mastered object clitics have mastered object NAs in their L3 Spanish. We speculate that the availability of object clitics facilitates the acquisition of object NAs. If this correlation is correct, given the claim that object clitics are hard to acquire, that subject and object NAs are not acquired simultaneously is expected. Many issues remain open for future research. One issue arises from the claim that it is (φ-)feature compositions that may be subject to cross-linguistic variation (e.g., Saito, 2007). Given this, it is natural that the L3 development is characterized with the features available to L3 learners at different stages. This raises an important question for the cumulative enhancement model. Since the nature of Japanese NAs is different from that of their Spanish counterparts (Oku, 1998; Saito, 2007), it is not immediately obvious how similar these two languages are in terms of feature compositions, despite appearances, and thus, it is not immediately apparent why the presence of Japanese NAs facilitates acquisition of Spanish NAs in their L3 Spanish, if feature compositions are the locus of the comparison in point (see the section on L3 acquisition models). This important issue needs investigating to clarify the nature of L1 transfer and L2 influence (see Yamada & Miyamoto, 2017).

References

Alamillo, A. R., & Schwenter, S. (2007). Null objects and neuter lo: A cross-dialectal variationist analysis. In J. Holmquist (ed.), *Selected proceedings of the third workshop on Spanish sociolinguistics* (pp. 113-121). Somerville, MA: Cascadilla Press.

Bardel, C. & Falk, Y. (2007). The role of the second language in third language acquisition: The case of Germanic syntax. *Second Language Research*, *23*, 459-484.

Campos, H. (1986) Indefinite object drop. *Linguistic Inquiry*, *17*(2), 354-359.

Castilla, A., & Pérez-Leroux, A. T. (2010) Omissions and substitutions in Spanish object clitics: developmental optionality as a property of the computational system. *Language Acquisition*, *17*(2), 2-25.

Duguine, M. (2013). Null arguments and linguistic variation: A minimalist analysis of pro-drop. Ph.D. dissertation, University of Basque Country UPV/EHU and Université de Nantes.

Flynn, S., Foley, C., & Vinnitskaya, I. (2004) The cumulative-enhancement model for language acquisition. Comparing adults' and children's patterns of development in first, second and third language acquisition. *International Journal of Multilingualism*, *1*, 3-17.

Miyamoto, Y., & Yamada, K. (2015) On null arguments and phi-features in SLA: A preliminary study. *Nanzan Linguistics*, *10*, 1-20.

Na Ranong, S., & Leung, Y-K I. (2009) Null objects in L1 Thai-L2 English-L3 Chinese: An empirical take on a theoretical problem. In Y-K. I. Leung (ed.), *Third language acquisition and universal grammar* (pp.162-191). Bristol, UK: Multilingual Matters.

Oku, S. (1998). LF copy analysis of Japanese null arguments, *CLS*, *34*, 299-314.

Rothman, J. (2010). On the typological economy of syntactic transfer: Word order andrelative clause attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics*, *48*, 245-274.

Saito,, M. (2007). Notes on East Asian argument ellipsis. *Language Research*, *43*, 203-227.

Yamada, K., & Y. Miyamoto (2017). On the interpretation of null arguments in L2 Japanese by European non-pro-drop and pro-drop language speakers. *Journal of European Second Language Association*, *1*, 73-89.

To cite this article: Miyamoto, Y., & Yamada, K. (2017). On the mixed nature of L3 Spanish grammar of L1 Japanese subjects with L2 English. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 193-198). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Japanese pitch accent in an English/Nupe/Hausa trilingual

Becky Muradás-Taylor
b.muradas-taylor@yorksj.ac.uk

York St John University

Abstract. Pitch accent, which distinguishes words in Standard Japanese, is difficult for speakers of English to acquire. This is likely to be because pitch in English does not have lexical function. However, alternative explanations could be insufficient Standard Japanese input or lack of explicit instruction. This paper reports on an English/Nupe/Hausa trilingual learner of Japanese who uses Standard Japanese pitch accent accurately, in spite of no residence in Japan or explicit instruction on pitch accent. Nupe and Hausa are tonal i.e. have lexical pitch. The aim of the paper is to report on the accuracy and stability of the participant's pitch accent; to consider how their language background has aided this acquisition, and to discuss implications for monolingual English speaking learners of Japanese. The data consists of a three minute audio recording of a presentation given in Japanese. The participant produced 90% of words with accurate Standard Japanese pitch accent and 93% of repeated words with accurate stable pitch accent. The participant's successful acquisition of pitch accent is argued to be because of the presence of lexical pitch in Nupe and Hausa. Since they achieved this despite no explicit instruction, stay in Japan, or a native-speaker tutor, the difficulty monolingual English speakers have acquiring pitch accent cannot be easily dismissed due to lack of explicit instruction or input. This finding provides support for the argument that English speakers' difficulty acquiring Japanese pitch accent is due to pitch not being lexical in English. However, other bilinguals (English + tone language) are needed to strengthen this claim.

Keywords: L2 acquisition, L1 influence, input, instruction, suprasegmental, stress, tone

Introduction

This is a case study of a person with three L1s – English and the Nigerian languages Nupe and Hausa – who started studying Japanese as an additional language in the UK at the age of 30. The study investigates the participant's production of pitch accent in their spoken Japanese, focussing on its accuracy (i.e. adherence to Standard Japanese norms) and stability (i.e. the extent to which repeated words have the same accent type). The participant produces pitch accent that is both accurate and stable. This is of interest because previous research on the acquisition of pitch accent by monolingual speakers of English learning Japanese showed that pitch accent was neither accurate nor stable (Taylor, 2012). Since the trilingual participant has acquired Standard Japanese pitch accent despite not receiving explicit instruction on pitch accent nor living in Japan, the difficulty that monolingual English speakers have acquiring pitch accent cannot be attributed to insufficient input or lack of explicit instruction. Instead the difference between the trilingual participant and the monolingual English speakers in Taylor (2012) is argued to be due to the trilingual participant's two other L1s, Nupe and Hausa, which are both tonal, i.e. have lexical pitch.

Pitch in Japanese, Nupe, Hausa and English

Pitch in Japanese is used to distinguish words. For example in Standard Japanese *hashi* 'chopsticks' has initial accent and *hashi* 'bridge' has final accent. What distinguishes different accent types is the presence or absence of an accent (realised phonetically as a sharp fall in pitch) and its position (Vance, 2008). Initially-accented words have an accent on the first syllable; finally-accented words have an accent on the final syllable; unaccented words have no accent i.e. no sharp pitch fall. Words longer than two syllables can have an accent on the second, third, fourth, etc. syllable: for simplicity, they are often referred to as having 'medial' accent.

Pitch in Nupe and Hausa is also lexical. Unlike Japanese pitch accent, where the presence or absence of an accent and its position distinguish the accent types off words, Nupe and Hausa are tonal, with syllables specified for tone. Nupe has five tones: high, mid, low, rise, fall; Hausa has three: high, low, fall (Yip, 2002).

Pitch in English, however, does not have lexical function. Instead, stressed syllables have high, low, falling or rising pitch (Pierrehumbert, 1980). The pitch shape differences depend on whether the utterance is, for example, a statement, question, or conveying surprise.

That pitch in English does not have lexical function is central to this study, but may go against the intuitions of users of English. Even in academic writing, pitch is sometimes described as an acoustic correlate of English stress. Fry (1958) is often quoted as evidence of this: for minimal pairs such as the noun *object* and the verb *object*, listeners used pitch to distinguish these words. However, calling pitch a correlate of stress is a ‘common misunderstanding’ (Beckman & Edwards, 1994:13) caused by a confound between lexical stress and utterance-level pitch (Sluijter & van Heuven, 1996a). Stressed syllables are potential ‘docking sites’ for utterance-level pitch (Sluijter and van Heuven, 1996b:2471); although listeners may use pitch to distinguish two words in citation form, it does not follow that pitch is a correlate of stress. Nor is pitch a reliable cue to stress: words that are not focussed do not have a pitch movement (Sluijter & van Heuven, 1996a, 1996b) and, for focussed words, the shape of the pitch movement depends on whether the utterance is a question, statement, conveying surprise, etc.

The following example illustrates the difference between pitch in Japanese (lexical function) and English (post-lexical function). If you take the word *happy* and produce it with rising intonation, you get the word *happy* as a question. However, if you take the initially-accented word *hashi* ‘chopsticks’ and produce it with rising pitch, you get a different word: finally accented *hashi* ‘bridge’. Since pitch has lexical function in Japanese, change the pitch of *hashi* and you change the word meaning. Since pitch does not have lexical function in English, change the pitch of *happy* (to falling pitch, rising pitch, falling-rising pitch, rising-falling pitch, etc.) and you will still have the same word.

Perception of Japanese pitch accent by L1 English speakers

The ability of L1 English-speaking L2 learners of Japanese to perceive Japanese pitch accent is task-dependent. In an AX discrimination task, where listeners judge whether two words have the same or different accent type, English-speaking learners perform similarly to L1 Japanese listeners (Hirano-Cook, 2011). English-speaking learners also perform similarly to L1 Japanese listeners on an ABX task, where listeners judge whether the word X has the same accent type as A or B (Sakamoto, 2011). On an identification task, however, where listeners identify which accent type a word has (e.g., initial, medial, final, unaccented), L1 Japanese speakers out-perform English-speaking learners (Hirano-Cook, 2011; Sakamoto, 2011). Similarly, unlike L1 Japanese speakers, English-speaking learners are unable to judge whether the accent types of words are correct (i.e. adhere to Standard Japanese norms; Shibata & Hurtig, 2008). Hirata (2015:736) concludes from findings such as these that L1 English-speaking L2 learners of Japanese can perceive pitch but do not categorise it like L1 Japanese speakers.

Interestingly, the correctness judgement task is difficult for all English-speaking learners, even advanced ones (Shibata & Hurtig, 2008). However, individual variation is observed on the identification task, with some learners performing well (Hirano-Cook, 2011). This may imply that English speaking learners of Japanese do not encode accent type into lexical representations in long-term memory, even if they learn to distinguish and identify them.

Production of Japanese pitch accent by L1 English speakers

Yamada (1994) reports on a learner who says the word *yappari* ‘obviously’ with three different accent types in one conversation. Such accent type instability is likely to be a result of pitch not having lexical function in English.

Taylor (2012) investigated pitch accent acquisition by English-speaking learners of Japanese, focussing on both accuracy and instability. This will be described in detail here, since it is the trilingual participant’s different behaviour to that of the participants in Taylor (2012) that is of interest in this study. The participants were L1 speakers of Standard Southern British English: a less-experienced group (n=13) who had studied Japanese for one or two years (mean 250 hours), and a more-experienced group (n=8) who had completed four years of a Japanese degree including a year in Japan. The learners read aloud 180 Japanese words in three contexts:

(i) In isolation e.g., *ame* 'rain'
(ii) Before a function word e.g., *ame da* 'it's rain'
(iii) Before a content word e.g., *ame ga furu* 'rain falls'

Japanese phoneticians identified the accent type of each word that the participants produced. The accuracy (i.e. percentage match with Standard Japanese) was 43% for the less-experienced group (lowest participant 32%, highest participant 52%, SD=6) and also 43% for the more-experienced group (lowest participant 36%, highest 48%, SD=5). The learners' accent types showed considerable instability: only 12% of the words produced by the less-experienced group, and 13% of the words produced by the more-experienced group had accent types that were produced accurately and stably across the three contexts.

The question that arises out of this research is why Japanese pitch accent is so difficult to acquire for English speakers. The most likely explanation is a linguistic one: because pitch is not lexical in English, L1 English-speakers do not encode pitch in their lexical representations when learning Japanese. This is supported by the perception findings described above: even if English-speaking learners learn to identify Standard Japanese accent types, they cannot say whether the accent types of words are correct. This implies that the learners' lexical knowledge not does contain information about pitch, even if accent types can be identified.

If pitch is not lexical in English, why do learners of Japanese not learn to encode it lexically on exposure to their L2? This could be because the functional load of pitch accent in Standard Japanese is low (Kitahara, 2001): pitch is rarely needed to disambiguate minimal pairs. Another possible reason is the considerable dialectal variation in Japanese, with words' accent types and even the accentual system varying between dialects (Kubozono, 2012). Together these could mean that L2 pitch accent is not necessary for learners and, therefore, not learned.

However, there could be other explanations. Perhaps the participants in Taylor (2012) had received insufficient Standard Japanese input (see e.g., Flege, 2009)? Or perhaps they had received insufficient explicit instruction (see e.g., Thomson and Derwing, 2015)? The second of these is particularly likely, since commonly used textbooks in the UK introduce pitch accent at the beginning, but do not refer to or practice it in later chapters (Shport, 2008).

Aims

The aims are: (i) to measure the accuracy and stability of the participant's pitch accent; (ii) to consider how the participant's language background has aided this acquisition; and (iii) to discuss implications for monolingual L1 English speaking learners such as those in Taylor (2012).

Method

Participant

The participant was raised in Nigeria and the UK. They (with 'they' being used as a singular gender neutral pronoun) were born in Nigeria, lived in the UK from age 1 to age 5, and Nigeria from age 6 to age 23. They were raised trilingually: English and the Nigerian languages Nupe and Hausa were all used in the home from birth. They were educated in English both in the UK and in Nigeria, and studied Hausa at school from age 9.

The participant started a beginner-level Japanese class at a UK university at the age of 30. The course was delivered by a British English tutor for two hours a week for 12 weeks. They then took the intermediate-level course with the same tutor which was three hours a week for a further 12 weeks. They did not receive any explicit pitch accent instruction on these courses. At the time that the data was collected, they had never been to Japan.

The key facts regarding the participant's language background are firstly that, as well as English, they have two further L1s, Nupe and Hausa, which are tonal and therefore have lexical pitch; and secondly, that the participant had not received explicit instruction on pitch accent and had not been to Japan.

However, the participant's language background is actually more complex: as well as their three L1s (English/Nupe/Hausa), they also learned Arabic from the age of 3 for Quranic recitation, and Hindi from the age of 6 through interaction with Hindi-speaking neighbours. They also spent time in Malaysia and India between the ages of 23 and 27. Additionally, before entering the beginner-level Japanese class, they had a long-standing interest in Japanese poetry, role-playing games and anime.

Data collection and analysis

The data was a three-minute audio recording of a presentation on a topic chosen by the participant. The presentation was part of an assessment; informed consent for it to be used for research was obtained subsequently.

The data was analysed by comparing the accent type that the learner used for each word against the Standard Japanese norm as described in the Shin Meikai accent dictionary (Kindaichi & Akinaga, 2001). This was used to give a percentage of words with accurate accent types. In addition, where words appeared more than once in the recording, their accent types were coded as the same or different, and a percentage of words with stable accent types calculated.

Unlike in Taylor (2012), where the participants read aloud a list of words and phrases, the first stage of analysis in this study involved deciding what to count as a word. The accent types of a 'content word + function word' unit were judged where possible, even if there was a pause before the function word. The second stage involved deciding whether to exclude any words due to their accents being deleted or compressed. In Standard Japanese, two accented words can be grouped together in one 'accentual' phrase, with the second accent deleted, or in one 'intonational' phrase, in which case the second accent is retained but compressed or 'downstepped' (Venditti, 2005:175). An attempt was made to exclude such contexts systematically but did not prove possible. Instead, words were included in the analysis, if it was possible to identify their accent type aurally.

In total, the accent types of 113 words were identified (tokens, not types, as the stability of the accent types was also of interest). Fifteen words appeared more than once.

Results

The participant produced 102 out of 113 words (i.e. 90%) with accurate Standard Japanese accent types. Figure 1 compares the accuracy of the trilingual participant (TRI) to the 21 monolingual English learners of Japanese in Taylor (2012), where LE is the less-experienced group (n=13) and ME is the more-experienced group (n=8).

From Figure 1, it is clear that the learners in Taylor (2012) all have low accuracy, with little variation between learners, and the trilingual participant has high accuracy.

In addition, of the 15 words that were uttered more than once during the recording, 14 (93%) had the same accent type on both occasions. This contrasts with the learners in Taylor (2012), where even the more-experienced learners only produced 13% with stable accurate accent types.

Although the focus of this paper is the accuracy and stability of the participant's pitch accent, it is worth pointing out that the eleven words whose accent type was not Standard varied in part of speech (noun, verb, etc.), syllable number and structure, Standard Japanese accent type, and produced accent type. No pattern (e.g., initially-accented words consistently being produced with final accent, for example) was observed.

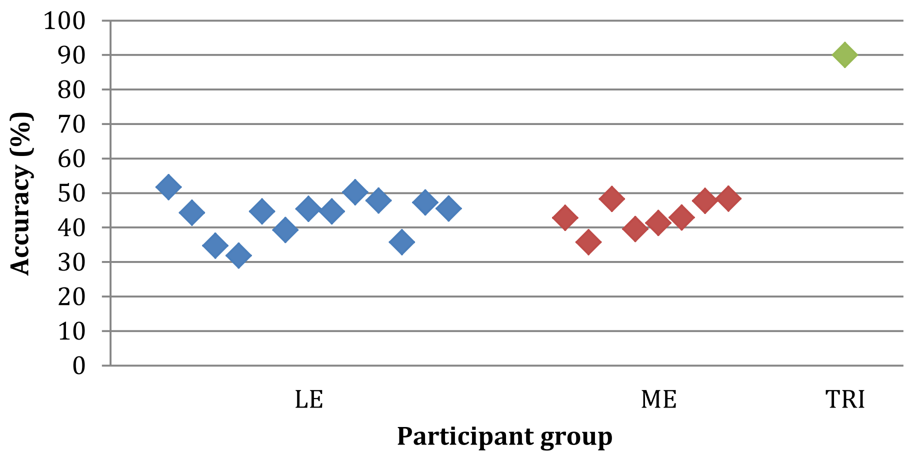


Figure 1. Accent type accuracy of the trilingual participant compared to the Taylor (2012) participants

Discussion

The participant's pitch accent was accurate and stable. 90% of words produced had accurate Standard Japanese accent types, and 93% of repeated words had accurate stable accent types. These findings are in striking contrast to the English monolingual speakers in Taylor (2012), where even the more-experienced group of learners only produced 43% of words with accurate Standard Japanese accent types and 13% with stable Standard Japanese accent types.

It was suggested earlier that the monolingual English speakers' difficulty acquiring Japanese pitch accent could be attributed to insufficient Standard Japanese input or lack of explicit instruction. However, the trilingual participant had only received 60 hours of Japanese language instruction over 24 weeks, had had a British English tutor, had had no explicit instruction on pitch accent, and had not lived in Japan. The participants in Taylor (2012) had received more Standard Japanese input (mean 250 hours for the less-experienced group; 970 hours for the more-experienced group) during their year in Japan. And neither the trilingual participant nor the Taylor (2012) participants had received explicit instruction on pitch accent. The difference between the trilingual participant and the monolingual English participants cannot therefore be attributed to these factors.

The most likely explanation for the difference between the trilingual participant and those in Taylor (2012) is the trilingual participant's two additional L1s with lexical pitch. These must have enabled the trilingual participant to encode Standard Japanese accent type into the lexical representation. Monolingual English speaking learners, however, must not encode pitch in the lexical representation, even if they are able to distinguish different accent types and identify them correctly.

Note that the trilingual participant produces Japanese pitch accent with high accuracy and stability, despite the fact that Japanese has pitch accent, with one or no syllables in a word specified as having a pitch fall, whereas Nupe and Hausa have tone with syllables specified as high, mid, low, rise, fall for Nupe, and high, low, fall for Hausa. Further research could investigate whether this difference has any effect on the participant's L2 Japanese.

A limitation of the current study is that the participant appears to be a particularly good language learner, who picked up Hindi from neighbours, for example. They (singular) also had considerable exposure to Japanese culture before starting formal study, casting doubt on the claim that their Standard Japanese input is limited. However, many of the monolingual English speakers in Taylor (2012) also expressed a keen interest in Japanese culture, and yet the highest accuracy of any learner was only 52%. This returns us to the original argument that is a linguistic factor (the presence or absence of a tone language L1) not amount of Standard Japanese input that explains the difference between the trilingual participant and the monolingual ones.

Conclusion

This paper reports on the pitch accent production of a learner of Japanese who has the tone languages Nupe and Hausa as L1s, as well as English. This participant's accurate and stable Standard Japanese pitch accent contrasts with monolingual L1 English-speaking learners of Japanese, whose pitch accent is neither accurate nor stable (Taylor, 2012). Since the trilingual participant had received no explicit instruction on pitch accent, less Japanese instruction than the Taylor (2012) participants, and had not lived in Japan, the monolingual participants' difficulty acquiring Standard Japanese pitch accent cannot be dismissed as due to lack of explicit instruction or insufficient Standard Japanese input.

Instead, the difference between the participants is argued to be due to the presence or absence of an L1 with lexical pitch. Since pitch in English has post-lexical function, monolingual English speaking learners of Japanese are not expected to encode pitch accent into long term memory. In contrast, the trilingual participant's tonal L1s Nupe and Hausa must enable the encoding of Japanese accent types into long term memory, resulting in stable and accurate Standard Japanese pitch accent production.

In order to strengthen this claim, other bilinguals (English + tone language) are needed. Further research could include acoustic analysis, as an aural impression of the current participant is that their accent types are not just phonologically accurate, but similar to Standard Japanese in phonetic realization, too.

References

Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (ed.), *Phonological structure and phonetic form: Papers in laboratory phonology III* (pp. 7-33).

Flege, J. (2009). Give input a chance! In T. Piske & M. Young-Scholten (eds), *Input Matters in SLA* (pp. 175-190).

Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech, 1*(2), 126-151.

Hirano-Cook, E. (2011). *Japanese pitch accent acquisition by learners of Japanese: Effects of training on Japanese accent instruction, perception, and production* . PhD Thesis, University of Kansas).

Hirata, Y. (2015). L2 Phonetics and phonology. In H. Kubozono eEd.), *Handbook of Japanese phonetics and phonology* (pp. 719-762). Berlin: De Gruyter Mouton.

Kindaichi, H., & Akinaga, K. (eds.) (2001). *Shin Meikai nihongo akusento jiten [New Meikai Japanese accent dictionary]*. Tokyo: Sanseido.

Kitahara, M. (2001). *Category structure and function of pitch accent in Tokyo Japanese*. PhD Thesis, Indiana University.

Kubozono, H. (2012). Varieties of pitch accent systems in Japanese. *Lingua*,*122*(13), 1395-1414.

Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation* (PhD Thesis, Massachusetts Institute of Technology.

Sakamoto, E. (2011). *Investigation of factors behind foreign accent in the L2 acquisition of Japanese lexical pitch accent by adult English speakers*. PhD Thesis, University of Edinburgh.

Shibata, T., & Hurtig, R. R. (2008). Prosody acquisition by Japanese learners. In Z. Han & E. S. Park (eds.), *Understanding second language process* (pp. 176-204). Clevedon, UK: Multilingual Matters.

Shport, I. A. (2008). Acquisition of Japanese pitch accent by American learners. In P. Heinrich & Y. Sugita (eds.), *Japanese as foreign language in the age of globalization* (pp. 165-187). München: Iudicium Verlag.

Sluijter, A. M. C., & van Heuven, V. J. (1996a). Acoustic correlates of linguistic stress and accent in Dutch and American English. *Proceedings of the Fourth International Conference on Spoken Language Processing*, *2*, 630-633.

Sluijter, A. M. C., & van Heuven, V. J. (1996b). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America, 100*(4), 2471-2485.

Taylor, R. L. (2012). *Eigo washa ni yoru nihongo no go akusento no shuutoku [The acquisition of Japanese lexical accent by English speakers]*. PhD Thesis, Nagoya University.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*,36, 326-344.

Yamada, N. (1994). Nihongo akusento shuutoku no ichidankai—gaikokujin gakushuusha no baai— [A stage in acquiring Japanese accent: the case of foreign learners]. *Journal of Japanese Language Teaching*, 83, 108-120.

Vance, T. (2008). *The sounds of Japanese.* Cambridge, UK: Cambridge University Press.

Venditti, J. J. (2005). The J_TOBI model of Japanese intonation. In S. Jun (ed.), *Prosodic typology and transcription: A unified approach* (pp. 172-200). Oxford, UK: Oxford University Press.

Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.

To cite this article: Muradás-Taylor, B. (2017). Japanese pitch accent in an English/Nupe/Hausa trilingual. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 199-205). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Dual-lingual interactions with passive bilingual children

Janice Nakamura
nakamura_janice@isc.sagami-wu.ac.jp

Sagami Women's University

Abstract. The inability or reluctance of bilingual children to speak one of their languages is a phenomenon that baffles researchers and parents. This study contributes to this relatively unexplored subject by examining how parental discourse strategies (Lanza, 1992, 1997) play a role in dual-lingual interactions with passive bilingual children. Naturalistic audio data was collected from two children who speak mainly Japanese to their Italian-speaking and English-speaking parents, respectively. The results revealed that both parents predominantly responded to their children's Japanese utterances using the move-on strategy. By continuing with the conversation, the parents did not encourage their children's production of their weaker language. Explicit discourse strategies (e.g., instruction to translate) that request for child target language production were rarely used. When employed, these strategies were not very effective in eliciting production. These findings suggest that while the prevalent use of move-on strategy by the parents perpetuated dual-lingual interactions, greater use of explicit discourse strategies may not activate the children's bilingualism. Other means of promoting weaker language production need to be explored.

Keywords: passive bilingualism, parental discourse strategies, dual-lingual

Introduction

Children who receive bilingual exposure do not necessarily produce both of their languages. In Japan, roughly one in three English-Japanese bilingual children speak only Japanese (Billings, 1990; Noguchi, 2001; Yamamoto, 2001). When a bilingual child does not produce one of his languages but his or her caregiver continues speaking it, both parties engage in dual-lingual interactions (Saville-Troike, 1987). These interactions are potentially problematic and may inhibit the harmonious development of two languages (De Houwer, 2015). They require increased agency and effort, and caregivers can get demoralized when children continually do not reciprocate in the same language that they are spoken to (Smith-Christmas, 2016).

Dual-lingual interactions may be attributed to parental use of discourse strategies. When the child uses the 'wrong' language, the parent can respond using six types of discourse strategies (Chevalier, 2013; Döpke, 1992; Lanza, 1992, 1997). While they are all termed as 'strategies' in this paper, some discourse strategies are undeliberated responses to children's language use. Instruction to Translate, Minimal Grasp, Expressed Guess, and Adult Repetition are explicit strategies that intentionally request for the child to speak the target language, whereas Move-On and Codeswitching strategies are implicit responses where the parent continues with the conversation without much reflection. Instruction to Translate is the most explicit strategy where the child is requested to rephrase a preceding utterance, e.g., *how do you say it in English?* Minimal Grasp is an explicit strategy where the parent feigns non-comprehension, e.g., *what did you say?* The parent can also make an Expressed guess, e.g., *do you want juice?* In the Adult Repetition strategy, the parent provides the correct language form, e.g., *I want juice*. Move-On is an implicit strategy where the parent continues with the conversation, e.g., *there's none left*. Codeswitching is another implicit strategy. Switching to the other language deprives the child of target language input.

In studies on active bilingual and trilingual children, the use of explicit discourse strategies (e.g., Instruction to Translate and Minimal Grasp) is linked to higher target language production. Contrastively, the use of implicit strategies (i.e., move-on and codeswitching) is associated with higher rates of language mixing (e.g., Chevalier, 2013; Juan-Garau & Pérez-Vidal, 2001). However, no research on discourse strategies has been conducted on passive bilingual children, and we cannot assume that discourse strategies that are effective on active bilinguals will work equally well on them. Given that passive bilingual children can become active users of their languages later (De Houwer, 2006; Quay 2001), we need to understand how parental discourse strategies contribute towards children's passive bilingualism and the dual-lingual interactions that they have with their parents.

Method

Two bilingual children who were born and raised in Tokyo, Japan were recruited for this study. While they received bilingual exposure from birth in the One-Parent-One-Language setting, Max (age 7) and Nina (age 4) both use Japanese predominantly with their Italian-speaking and English-speaking fathers, respectively. Their fathers understand and speak Japanese, but speak their own languages to their children. Each father-child pair made audio recordings at home while they were playing games, looking at books or doing homework. Six audio recordings for a total of 4.5 hours from Max and his father, and eight audio recordings totalling 3 hours from Nina and her father were transcribed by two research assistants who were native or fluent speakers of Italian, English, and Japanese. The transcripts were coded using CHAT and analyzed using CLAN (MacWhinney, 2000). All parent and child utterances were coded as Japanese, English, Italian or mixed. Utterances that were unintelligible or did not belong to either language were excluded (e.g., interjections). Parental utterances that immediately followed the children's Japanese utterances were also coded for the type of discourse strategy used.

Results

Language use

Figure 1 shows the percentage of utterances produced by the parents and children in each language. The results indicate that parent-child utterances were largely but not completely dual-lingual. Max's father's utterances were 97.3% Italian (ITA), but only 19.7% of Max's utterances were in the same language. Nina's father's utterances were 98.6% English (ENG), but only 40.3% of Nina's utterances were English. Japanese utterances made up 73.9% and 55.8% of Max's and Nina's total utterances respectively. Only a small proportion of utterances were mixed (MIX) for all speakers. While both children used Japanese predominantly, they did speak some Italian or English, contrary to initial reports by their parents that they were passive bilinguals. Nevertheless, their target language use was highly limited. In particular, Nina often produced repetitive, rote-learned English utterances, of which many contained only single words. Her English use is illustrated in extract 2.

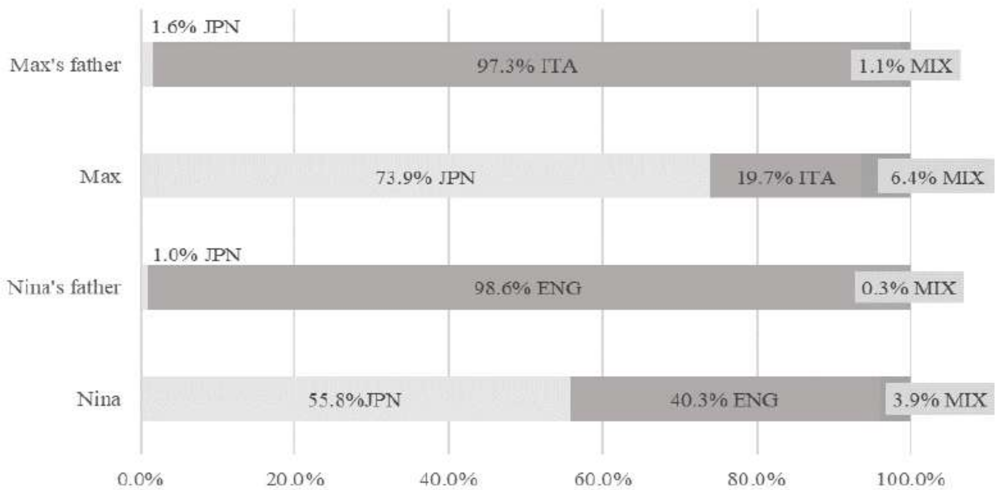


Figure 1: Percentage of utterances produced in each language

Discourse strategies

Table 1 summarizes the use of discourse strategies according to type. The results were similar for both parents. The move-on strategy was used about 93% of the time. Other strategies were seldom used. Both parents codeswitched in response to the children's Japanese utterances only around 4% of the time. The parents' Japanese utterances were usually once-off repetitions of the children's preceding Japanese utterances. Both parents reverted to their own languages immediately after producing a Japanese utterance, so there was no risk of them shifting to Japanese. Their lack of codeswitching also maintained the dual-lingual nature of their interactions.

Table 1. Parents' use of discourse strategies in response to children's Japanese utterances

		IT	MG	EX	AR	MV	CS	Total
Max's father	No. of times used	6	0	0	6	393	16	421
	Percentage	(1.4%)	(-)	(-)	(1.4%)	(93.4%)	(3.8%)	(100%)
Nina's father	No. of times used	0	0	6	4	390	18	418
	Percentage	(-)	(-)	(1.4%)	(1.0%)	(93.3%)	(4.3%)	(100%)

Notes on abbreviations: instruction to translate (IT), minimal grasp (MG), expressed guess (EX), adult repetition (AR), move on (MV), and codeswitching (CS).

Extract 1 illustrates the continuous use of the move-on strategy by Max's father (FAT) to Max (MAX) when they were choosing a tree from a catalog to plant in their garden. The main tier (marked with an asterisk) represents the utterance whereas the secondary tier (marked with %com) shows the English translation. In line 1, Max asked in Japanese about where the tree should be planted. His father used the move-on strategy by telling him in Italian where to look (line 2). Max asked another question in Japanese to clarify (line 3), which his father replied in Italian (line 4). Max's third question (line 5) was also explained in Italian (line 6). The three consecutive uses of the move-on strategy created a perfectly dual-lingual parent-child interaction.

Extract 1

1. *MAX: Doko ni uemasu?
   %com: Where will we plant it?
2. *FAT: Devi guardare sopratutto questo.
   %com: This is where you have to look
3. *MAX: Ue?
   %com: up?
4. *FAT: Si perché noi qui abbiamo cosi poco spazio.
   %com: Yes because we only have a little space here
5. *MAX: Koko tate kara?
   %com: Because it is vertical?
6. *FAT: Ah in alto.
   %com: It is high

The move-on strategy was also prevalent in Nina's father's interactions with the child. Extract 2 illustrates their dual-lingual interactions while playing Jenga, a tower building game. Nina (NIN) urged her father (FAT) in Japanese to remove a block (line 1). Nina's father used the move-on strategy, by responding in English (line 2). When Nina expressed her concern in Japanese that the tower might break (line 5), her father responded in English again using the same strategy (line 6) and teased her that he was going to win the game (line 7). This made Nina protest in English by exclaiming *no* and *me* (italicized in lines 8 and 9). When she reverted to Japanese (line 13), the move-on strategy was used again (line 14). While this extract is not completely dual-lingual, it illustrates her father's consecutive uses of the move-on strategy. It also shows how Nina's typical English productions, which were often single-word utterances used for emphasis (lines 8 and 9), contrast with her longer and more complex Japanese utterances (lines 1, 5 and 13).

Extract 2

1. *NIN: papa mo ikko yatte ii yo.
   %com: papa can do one more.
2. *FAT: okay.
3. *FAT: alright.
4. *FAT: thanks for letting me join in this game today.
5. *NIN: papa kowarechau kamoshirenai.
   %com: papa it might break
6. *FAT: you think?
7. *FAT: I think I am going to win.
8. *NIN: <*no*> [/] *no*!
9. *NIN: *me*!
10. *FAT: really?
11. *FAT: well this should be interesting.
12. *FAT: you know I don't think I ever played against you in this game.
13. *NIN: kore ne itsumo mama to yatteiru kara yatta koto ga aru da yo.
    %com: I always play this with mummy so I have played it before.
14. *FAT: I know.

Explicit strategies are generally considered effective for eliciting child target language production. However, in this study, they were used even more rarely than codeswitching strategies. Max’s father used instruction to translate and adult repetition strategies occasionally (both 1.4% of total discourse strategies). Likewise, Nina’s father only used the expressed guess and adult repetition strategies several times (1.4% and 1.0% of total discourse strategies respectively). The minimal grasp strategy was not used at all by both parents. Further analysis was conducted to determine the extent to which explicit strategies managed to elicited child target language production.

Table 2. Children’s production in response to parents’ use of explicit discourse strategies

		Instruction to translate	Minimal grasp	Expressed guess	Adult repetition
Max’s father	No. of times used	6	-	-	6
	Child production	3	-	-	0
Nina’s father	No. of times used	-	-	6	4
	Child production	-	-	0	1

Table 2 shows that the explicit discourse strategies were not effective on the children. Except for one English utterance from Nina, the expressed guess and adult repetition strategies elicited no target language production. However, Max’s father’s use of the instruction to translate strategy managed to elicit Italian three out of six times that it was used. Extract 3 shows how the instruction to translate strategy was used successfully. In this extract, Max said in Japanese that he would draw a picture of an egg (line 1). When his father instructed him to translate *tamago*, the Japanese term for ‘egg’, into Italian (line 2), the child complied by producing the Italian equivalent *ouvo* (italicized in line 3). While not shown extract 3, subsequent prodding by his father also led to Max’s production of Italian terms such as *scrivere* ‘write’, *disegno* ‘picture’ and *disegnare* ‘the act of drawing’.

Extract 3

1. *MAX: Tamago no e wo egaku.
   %com: I will draw a picture of an egg
2. *FAT: Tamago si dice.
   %com: How do you say tamago?
3. *MAX: *Uovo*.
   %com: Egg

While the instruction to translate strategy had some success, the adult repetition strategy did not encourage Max to produce Italian at all. In extract 4, Max asked for the calculator in Japanese (line 1), his father instructed him to translate the Japanese term, *keisanki* ‘calculator’ into Italian (Line 2). Max did not comply with this request and said *blah blah* (line 3). The father then used the adult repetition strategy by prompting Max with the first syllable from the Italian equivalent, *calcolatrice* (line 4). However, Max just responded with an unintelligible sound (line 5). Max did not respond favorably to the two strategies probably because he was more interested in finding the calculator to calculate scores in a game.

Extract 4

1. *MAX: Keisanki wa?
   %com: Where is the calculator?
2. *FAT: Come si dice keisanki?
   %com: How do you say keisanki?
3. *MAX: <blah> [/] blah.
4. *FAT: Si dice cal.
   %com: It is cal.
5. *MAX: x.
6. *FAT: No.
   %com: No

Likewise, Nina did not respond to the adult repetition strategy. In extract 5, Nina produced the mixed utterance *Jaja is make* (line 1). *Jaja* referred to her father, who rephrased Nina’s utterance in English as *Jaja lost* (line 2). However, he did not pause but continued talking (lines 3, 4 and 5). Therefore, Nina was not cued to repeat her Japanese utterance in English. While her father modeled the right language form for her, he did not seem to expect her to reproduce it.

Extract 5

1. *NIN: Jaja is make.
   %com: Jaja lost
2. *FAT: Jaja lost.
3. *FAT: I can imagine why.
4. *FAT: miss spinderella spins a six.
5. *FAT: everything you spin is a six.

Extract 6 was the only instance where the adult repetition strategy led to child production. In this extract, Nina could not produce the word *ladybug* in both of her languages. However, when her father prodded her, Nina could produce the Japanese equivalent, *tentomushi* (line 5). Her father then used the adult repetition strategy by providing her with the correct English term in line 6. The father did not continue with the conversation right after this strategy was used and waited for Nina to reproduce the word *ladybug*, which she did successfully (italicized in line 7).

Extract 6.

1. *FAT: what is that?
2. *NIN: wakanai.
   %com: I don’t know
3. *FAT: what's that in Japanese?
4. *FAT: you know what that is don't you?
5. *NIN: tentomushi.
   %com: ladybug
6. *FAT: in English we say a ladybug.
7. *NIN: *ladybug*.
8. *FAT: yep.

Discussion

While the Italian and American parents in this study used their own languages consistently, their children mainly spoke Japanese to them. Analysis of discourse strategies revealed that both parents predominantly used the move-on discourse strategy in response to their children's Japanese utterances. The prevalent use of the move-on strategy seemed to have perpetuated the dual-lingual nature of parent-child interactions and revealed a parental interactional style which focused on continuing the conversation instead of aligning their children's languages with their own. Explicit strategies were not often used probably because the parents felt that their children did not have the productive ability to rephrase many of their Japanese utterances in their weaker language, or would be reluctant to do so. It may also be possible that, in demonstrating verbal parent-child interaction in the audio recordings, the parents may have used fewer explicit strategies than they normally do. Nevertheless, explicit strategies were probably not used very frequently in daily interactions given the high production of Japanese by the children. These strategies are time-consuming because they require a conversation to be halted temporarily until the children rephrase a preceding Japanese utterance in their weaker languages.

Whenever explicit discourse strategies were used, they were not very successful in eliciting target language production. For them to work, the child must not only know the language form in his weaker language, but he must also be willing to produce it. Ill-timed use of explicit discourse strategies are not likely to be effective as demonstrated in Max's refusal to translate or repeat the Italian equivalent of the term calculator in extract 4 because he was more concerned about looking for the calculator. These strategies may only elicit target language production when the child is more compliant (cf. extract 3).

The expressed guess and adult repetition strategies were equally ineffective in making Nina rephrase her Japanese utterances in English. However, unlike Max, who chose not to oblige with such requests, she did not use English because her father did not pause for her to respond. As extract 5 shows, despite providing the correct language form for Nina, her father continued speaking without waiting for her to repeat it. Given her young age, Nina may be more willing than Max to respond to an adult repetition strategy. Therefore, Nina's father may have more success if he anticipates, and waits for her English production.

Nevertheless, any success explicit discourse strategies have on eliciting target language is likely to be limited because single word productions (e.g., *uovo* 'egg' or *ladybug*) are inadequate for carrying on a conversation in the target language. While they may elicit a response, they do not encourage spontaneous language use, which is essential for children to become active users of the language. These findings suggest that, while parental discourse strategies may have contributed to dual-lingual interactions with the child, they may not be useful in reversing this trend because the children lacked productive ability, and were aware that their parents understood their Japanese. It may better to give the children more exposure to monolingual speakers of the language such as grandparents in the home country to encourage production. The results of this study also suggest that parents need to be reflective of their language practices from the very beginning, i.e., at the early stages of the child's language development, and aware of the extent that their children are producing the target language to prevent dual-lingualism from arising. When parents are not vigilant about children's language use and continuously adopt the move-on strategy in interaction, passive bilingualism becomes a likely risk.

Conclusion

The inability or reluctance of some bilingual children to speak one of their languages is a phenomenon that baffles researchers and parents. The present study on the role of parental discourse strategies in dual-lingual interactions with two passive bilingual children is only the first step in understanding this phenomenon. The results revealed that both parents predominantly used the move-on strategy, which perpetuated the dual-lingual nature of parent-child interactions. By continuing with the conversation, the parents did not encourage their children's production of their weaker language.

Explicit discourse strategies were rarely used, and whenever they were employed, they were not successful in eliciting target language production. Greater use of explicit discourse strategies may not activate the children's bilingualism given their lack of productive ability. Other means of encouraging passive bilingual child's production in the weaker language need to be explored in future research.

Acknowledgements

This work was supported by the Japan Society for the Promotion of Science under the Grant-In-Aid for Young Scientists (B) (16K16867).

References

Billings, M. L. (1990). Some factors affecting the bilingual development of bicultural children in Japan. *AFW Journal*, April, 93-108.

Chevalier, S. (2013). Caregiver responses to language mixing by a trilingual. *Multilingua*, *32*(1), 1-32.

De Houwer, A. (2006), Bilingual language development: Early years. In K. Brown (ed.), *Encyclopedia of language and linguistics* (pp. 780-786). Oxford, UK: Elsevier.

De Houwer, A. (2015). Harmonious bilingual development: Young families' well-being in language contact situations. *International Journal of Bilingualism*, *19*(2), 169-184.

Döpke, S. (1992). *One parent one language: An interactional approach*. Amsterdam, Netherlands: John Benjamins.

Juan-Garau, M., & Pérez-Vidal, C. (2001). Mixing and pragmatic parental strategies in early bilingual acquisition. *Journal of Child Language*, *28*(1), 59-86.

Lanza, E. (1992). Can bilingual two-year-olds code-switch? *Journal of Child Language*, *19*(3), 633-658.

Lanza, E. (1997). *Language mixing in infant bilingualism: A sociolinguistic perspective*. Oxford, UK: Oxford University Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum.

Noguchi, M. G. (2001). Bilinguality and bicultural children in Japan: A pilot survey of factors linked to active English-Japanese bilingualism. In M. G. Noguchi & S. Fotos (eds.), *Studies in Japanese bilingualism* (pp. 234-271). Clevedon, UK: Multilingual Matters.

Quay, S. (2001). Managing linguistic boundaries in early trilingual development. In J. Cenoz & F. Genesee (eds.), *Trends in bilingual acquisition* (pp. 149-199). Amsterdam, Netherlands: John Benjamins.

Saville-Troike, M. (1987). Dilingual discourse: The negotiation of meaning without a common code. *Linguistics*, *25*(1), 81-106.

Smith-Christmas, C. (2016). *Family language policy: Maintaining an endangered language in the home*. Basingstoke, UK: Palgrave MacMillan.

Yamamoto, M. (2001). *Language use in interlingual families: A Japanese-English sociolinguistic study*. Clevedon, UK: Multilingual Matters.

To cite this article: Nakamura, J. (2017). Dual-lingual interactions with passive bilingual children. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 206-212). ISBN: 978-618-82351-1-3 URL: http://ismbs.eu/publications-2017

Pathbreaking verbs for developing recursion in the speech of children acquiring English

Anat Ninio
anat.ninio@huji.ac.il

The Hebrew University of Jerusalem

Abstract. The syntactic structure of sentences is represented by a tree which is built by the recursively applied operation Dependency (or Merge). Hauser, Chomsky, and Fitch (2002) claim recursion is innate and not learned. The alternative hypothesis is that children come to understand the principle of recursion on the basis of the linguistic input that transparently demonstrates the iterative nature of syntactic connection. We tested the evidence for processes of learning. Previous studies found that children first learn to use syntactic constructions with certain very frequent and prototypical pathbreaker verbs. We are now asking, are there pathbreaking verbs for recursion in children's early multiword utterances? We collected a large corpus of sentences of three words or longer possessing syntactic structure produced by 409 English-speaking young children from CHILDES (MacWhinney, 2000). We identified the word organizing each sentence, namely, the root of the syntactic tree. Most roots were verbs. There were 16 children who used only one single verb-lemma as root, and we consider these children beginners in the use of recursion. Just two verbs served for the great majority of sentences, *be* (44%) and *want* (38%). These two verbs were also used by most children who employed several different verbs as roots. Recursion has two pathbreaking verbs in child speech, *want* and *be*. Next, we analyzed three- and four-word long sentences of 508 parents from CHILDES, assuming that children may not be able to parse longer sentences and use them as their initial input for learning the principle of recursion. Close to 70% of them were built on some form of *be* as the root, but only a few (1.4%) were built on the verb *want*. The grammatical verb *be* requires two or more complements; it rarely if ever occurs without involving recursion. The results support the hypothesis that children learn the principle of recursion from the most frequent, prototypical, and generic exemplars in the linguistic input, apparently from sentences with the pathbreaking verb *be.*

Keywords: syntactic development, recursion, pathbreaking verbs, input, grammatical verbs

Introduction

This study belongs to a line of research that asks the question, how do children master syntactic structure-building? The first thing to emphasize is that although the present study represents an empirical learning approach to syntactic development, this is not a 'constructivist' or a 'usage-based' account of acquisition. We embrace a formal linguistic theoretical framework which we believe children gradually acquire, and do not hold with the idea that children store a large number of unanalyzed sentences until they submit them to statistical distributional analysis that discovers ad-hoc syntactic categories and rules. We hold that formal linguistic principles, categories and rules provide the best descriptions of people's way of constructing sentences. We believe children learn syntax from optimal exemplars which are the simplest, the most transparent, and ubiquitously encountered items of input exemplifying the relevant linguistic units and rules. Optimal exemplar learning is a general process of development found in cognitive systems as far from syntax as the immune system (Hershberg & Ninio, 2004). Potentially, such learning is one-shot, if the circumstances are optimal (see Nelson, 1987, on rare event learning). Thus, our empiricist model posits that learning is efficient and low-cost, and that our task as researchers is to find the environmental affordances for it, rather than assume that either innate principles or massive statistical analysis are involved in the process.

Our starting point is the linguistic computational system that builds sentences, the abstract core of language that consists of computational operations. We hold that this system uses a recursive combining operation called Dependency (Hudson, 1990). In the Minimalist Program a mathematically equivalent operation is called Merge (Chomsky, 1995).

The posited linguistic computational system follows the principles by which multi-element systems are constructed in computer science and mathematics. Taking syntax as the system of lawful composition of large structures from a set of combining units, it appears that a generative syntax needs just two elements: a base operation of connectivity and a process of recursion. This is a basic tenet of computer science (Abelson, Sussman, & Sussman, 1996).

The syntactic structure of sentences is represented by a dependency tree (Figure 1). The arrow leads from the head to the dependent; the head determines the character of the combination.

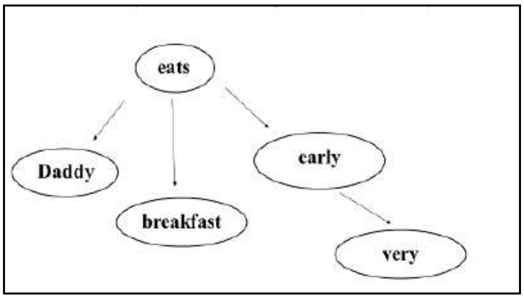


Figure 1. The syntactic structure of the sentence '*Daddy eats breakfast very early'*

Trees in mathematics are structures which are built recursively (Epp, 1995). This means the operation Dependency is applied repeatedly to connect more and more word-couples.

The base case of tree-building is a single couplet of word-nodes connected by the operation Dependency (Figure 2):

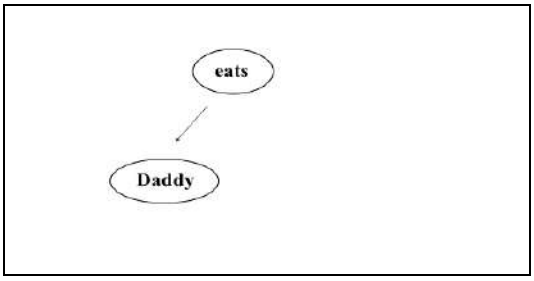


Figure 2. The syntactic structure of the combination '*Daddy eats*'

One can repeat the operation by giving one of the two existing nodes a Dependent (Figure 3). Applying the same combining operation to one of the nodes – that's the computational mechanism of recursion. Hauser, Chomsky, and Fitch (2002) claim that recursion is part of an innate Universal Grammar. The alternative hypothesis is that all syntactic principles, including recursion, are learned from the linguistic input. In previous studies, it was shown that children master the principle of dependency by learning it from transparent input sentences (Ninio, 2014a, b). Now we are asking how do children master recursion? We shall look for evidence of similar item-specific learning for recursion, from transparent input sentences. Basically any grammatical sentence three words long and longer involves the recursion of the dependency operation, by definition. When there is some reasonable doubt about the productivity of some types of multiword sentences, it is addressed explicitly.

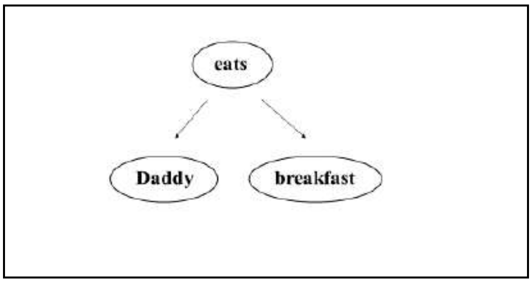


Figure 3. The syntactic structure of the sentence '*Daddy eats breakfast*'

Is recursion learned?

There are two processes of learning relevant to a novel structural principle and the actual constructions it is realized in: First, in a learning system, a new principle spreads gradually, with some first items leading the way and other items following after the while. In previous studies, it was found that children first learn to produce syntactic constructions with certain prototypical and generic verbs, which then serve as pathbreakers for other verbs (Goldberg, 2006; Ninio, 1999). Such leading verbs were found for the combinations verb-object, subject-verb-object and the like. Now we are assuming that the same phenomenon exists for the spread of novel structural *principles*. Pathbreaking verbs for verb-argument structures are not simply frequent items (as some authors believe, e.g., Theakston, Lieven, Pine, & Rowland, 2004) but, rather, they are generic exemplars that embody the relevant structural principle in a transparent way. We expect that the principle of recursion will be first expressed by children with such verb-headed sentences.

Second, in a learning system, structural principles are modeled in the input by frequent and transparent exemplars. These are usually the pathbreaking verbs children use to start the expression of the relevant principle, but even when they are not, the input affords the learning of the novel principle from highly frequent and transparent instances. In previous studies we found support for the idea that the input models novel structural principles such as the combining operation of dependency with transparent exemplars, in the shortest possible sentences (Ninio, 2014a, 2016).

We are now asking, are there pathbreaking verbs for recursion in children's early multiword utterances? Does linguistic input model transparent exemplars with high frequency? These verb-exemplars should enable, nay, force the children to acknowledge the recursive nature of the dependency operation. This means potential pathbreaking verbs should not be present in the input or in the output in sentences without recursion, *i.e.*, in two-word long sentences. This is a radical conclusion as developmental theories insist longer constructions are learned by combination or extension of children's own shorter ones (Elbers, 2000; Lieven, Salomo, & Tomasello, 2009).

Our hypotheses are:
1. Novel structural principles are first expressed by pathbreaking exemplars;
2. Structural principles are modeled in the input by frequent and transparent exemplars.

Method

Participants: children's corpus

English-language child samples were taken from the CHILDES (Child Language Data Exchange System) archive (MacWhinney, 2000), which is a public domain database for corpora on first language acquisition. All participants were observed in naturalistic, dyadic interaction of children and their parents. The observations were of normally developing young children with no diagnosed hearing or speech problems, and of their parents, native speakers of English. We restricted the contribution of each individual child to 300 spontaneous multiword sentences, starting from the first observation in which they produced multiword utterances. We collected a large corpus of sentences of three words or longer possessing syntactic structure produced by 409 children. The mean age of the children was 2;03,23 (*SD* = 4 months; range 14-42 months). We follow the tradition of researchers who examine pooled corpora of child speech for various characteristics thought to reflect on the relevant class of child speakers (Serratrice, Joseph, & Conti-Ramsden, 2003).

Results

Children's verbs at the root of multiword sentences involving recursion

In a learning system, a new principle spreads gradually, with some generic or prototypical first items leading the way. We are looking at evidence for the presence of pathbreaking verbs in children's sentences involving recursion.

We identified the word organizing each sentence, namely the root of the syntactic tree. The roots of the syntactic trees were mostly verbs. Children used 244 different verbs. The 10 most frequently used verbs were: *be, want, do, go, get, put, have, can, let,* and *see.*

Identifying beginners using a single-verb

Next, we grouped the child sample according to the number of different verbs used by the child as roots. There were 16 children who used only one single verb-lemma as root, and we consider these children beginners in the use of recursion. In this group, just two verbs served for the great majority of sentences, *be* (44% of the children) and *want* (38%). Examples 1a-j are the sentence types produced by the beginner-children with *be* as root.

1 a. He’s asleep.
b. It’s a kite.
c. My name is Gail.
d. That's my shoes.
e. There’s the ball.
f. What is that?
g. What’s this?
h. What's that?
i. Where’s the ball?
j. Because they’re yucky.

Examples 2a-j are the sentence types produced by the beginner-children with *want* as root.

2 a. I wanna put it back in there.
b. I want go bye bye.
c. I want leave.
d. I want snack.
e. I want that .
f. I want this.
g. I want to hold it.
h. I wanna put back in there!
i. I wanna put it back.
j. I want to put it back in there!
k. Want the big one.

Figure 4 presents the probability that children produce sentences with *be* and *want*, by the number of different roots they use, which is an estimate of the developmental phase the child is in.

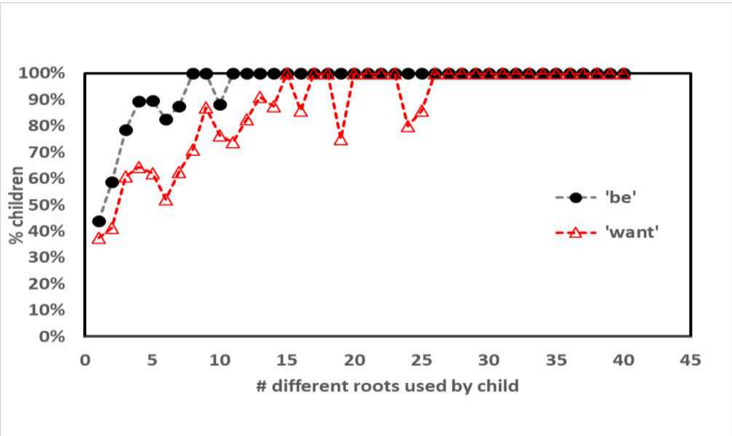


Figure 4. Probability that children produce sentences with *be* and *want*, by number of different roots they use (N=409)

The verbs *be* and *want* were also used by most children who employed two or more different verbs in their multiword sentences. The probabilities reach a plateau at 100% when the children use 10-15 different verbs as roots of longer sentences. Children use at least one of these verbs by the time they generate sentences with 3 verbs.

For comparison, the figure is repeated with the addition of the non-pathbreaking lexical verb *make* (Figure 5).

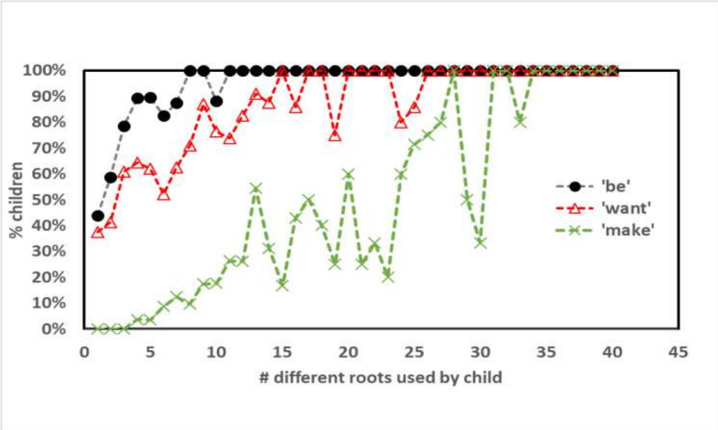


Figure 5. Probability that children produce sentences with *be, want* and *make*

A non-pathbreaking verb like *make* starts to be used as the root of sentences with recursion at a later developmental stage than *be* and *want*. Its use pattern is erratic, and reaches a plateau of 100% only when children produce 35 different verbs in such sentences. Similar patterns were observed with all other verbs serving as roots. We might conclude that recursion has two pathbreaking verbs in child speech, *want* and *be*, and other verbs.

The two pathbreaking verbs are rather different: *want* is a transitive lexical verb, with semantic arguments expressed as SVO; *be* is a grammatical verb with a subject and a second dependent which is a nonfinite verb or a predicate complement. Both dependents are usually considered not to be the verb *be*'s semantic arguments. It appears that recursion has two different prototypes, expressed in two differing syntactic constructions. Each may serve as the pathbreaker to a different subset of verbs: lexical verbs on the one side and grammatical verbs on the other.

As for syntactic valency, both verbs have a very strong requirement for two dependents, and as an indication, they rarely appear as the heads of two-word utterances where they receive only a single dependent. We checked the frequencies in children's and parents' two-word sentences, in the same samples used for the present study. Children produced 12,491 two-word sentences; among them only 1.33% was headed by *be*, and only 2.07% was headed by *want*. In the parental corpus, there were 23,204 two-word sentences; among them only 5.32% was headed by *be*, and only 0.44% was headed by *want*. The absence of such verbs from two-word long input and output sentences, demonstrates that *be* and *want* are exemplary two-dependent verbs. This characteristic turns them into prototypical roots for recursive structures.

One potential complication in considering *be* as a pathbreaking verb for recursion is that very often it is expressed in a contracted form, or cliticized on the subject. As an example we could take 1a., *he's asleep* as well as other early sentences with *be*. We should discuss the possibility sometimes raised in the literature that children learn the contracted combinations as a frozen single-word unit. In this case, the sentence including such forms does not involve recursion as it only consists of two words, for example in 1a. these are *he's* and *asleep*. If *He's* consists of a single word, there's only one dependency relation between it and the word *asleep*.

The counter-argument to this claim is that in this very developmental period, children express the same copular and temporal structures with the verb be omitted. For example, a sentence such as *He's asleep* would be expressed as *He asleep.* So called telegraphic combinations are direct evidence that children separate the subject and the copula as two words, one of which can be omitted. We checked the incidence of telegraphic combinations by the sample and found that 309 of the 409 children of the sample, that is, 75.55%, produced telegraphic sentences with omitted copulae and auxiliary verbs. In addition, the subject is sometimes omitted as well, leaving the grammatical verb and its complement as the single expressed terms. Such sentences often alternate in adjacent turns with the fully expressed construction including the copula or auxiliary verb and its two dependents. It appears children are aware of the complex structure of cliticised and contracted forms rather that seeing them as single words. We are justified as treating *be* as a verb involving recursion when it receives a subject and a complement, whatever way it connects with the subject expression.

Parents' verbs at the root of multiword sentences involving recursion

The second part of our research plan involves exploring input sentences that embody recursion, checking if they contain frequent, prototypical, and generic exemplars from which children can learn the structural principle of recursion.

Parents' corpus

English-language parental samples were taken from the same corpora in the CHILDES archive (MacWhinney, 2000) as the child sample. There were 508 English-speaking parents in the sample. The number of utterances included from each parent was restricted to a maximum of 3,000 addressed to the child, counting from the beginning of observations. We focused on parents' three-word and four-word long sentences possessing syntactic structure, excluding vocatives or interjections. Unfinished or cut off utterances, or containing words not transcribed were also excluded.

The parents' corpus consisted of 57,820 sentences, produced by 508 English-speaking parents addressing young children. This corpus represents the 'three-and four-word long' linguistic input that young children receive when acquiring the principle of recursion. The use of pooled corpora of unrelated parents as a representation of the linguistic input is a relatively conventional move in child language research (e.g., Goodman, Dale, & Li, 2008). Table 1 presents the ten most frequent verbs used as roots for parents' three- and four-word long sentences.

Table 1. Ten most frequent verbs used as roots for parents' three- and four-words long sentences (N=57,820)

Verb	Tokens	Percent %
be	39,788	68.81
do	5,506	9.52
can	2,465	4.26
go	2,161	3.74
have	1,106	1.91
want	807	1.40
get	592	1.02
like	557	0.96
know	359	0.62
come	298	0.51

Close to 70% of the sentences were built on some form of *be* as the root, but only a few (1.4%) were built on the verb *want*. The shortest and easiest input for recursion includes a very large proportion of grammatical verbs, not lexical verbs.

Discussion

In children's speech, there are two pathbreaking verbs for producing recursive structures: *be* and *want*. Of the two pathbreaking verbs, only *be* occurs with high frequency in parental sentences three or four words long, with easily parse-able structure. Such short sentences with *be* are probably the source of the principle of recursion in general. *Want* is rare in short parental sentences; probably such sentences are probably not the input for child sentences with *want* involving recursion. Recursion in sentences with lexical verbs as roots including *want* are probably the product of applying an already mastered principle – rather than being learned in a lexical-specific manner from the linguistic input.

References

Abelson, H., Sussman, G. J., & Sussman, J. (1996). *Structure and interpretation of computer programs* (2nd ed.). Cambridge, MA: The MIT Press/ McGraw-Hill.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Elbers, L. (2000). An output-as-input hypothesis in language acquisition. In P. Broeder & J. Murre (eds.), *Models of language acquisition* (pp. 244-271). Oxford, UK: Oxford University Press.

Epp, S. (1995). *Discrete mathematics with applications* (2nd ed.). Boston, MA: PWS Publishing Company.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford, UK: Oxford University Press.

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language, 35*, 515-531.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569-1579.

Hershberg, U., & Ninio, A. (2004). Optimal exemplar learning in cognitive systems. Invited article for a special issue of *Cognitive Systems (ESSCS)* on Multidisciplinary Aspects of Learning, edited by G. Dalenoort & M. Pichat, *6*, 181-188.

Hudson, R. (1990). *English word grammar*. Oxford, UK: Basil Blackwell.

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics*, *20*, 481-508.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.

Nelson, K. E. (1987). Some observations from the perspective of the rare event cognitive comparison theory of language acquisition. In K. E. Nelson & A. Van Kleeck (eds), *Children's language* (Vol. 6, pp. 289-331). Hillsdale, NJ: Lawrence Erlbaum.

Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, *26*, 619-653.

Ninio, A. (2014a). Learning a generative syntax from transparent syntactic atoms in the linguistic input. *Journal of Child Language*, *41*, 1249-1275.

Ninio, A. (2014b). Variables and values in children's early word-combinations. *Psychology of Language and Communication*, *18*(2), 106-125.

Ninio, A. (2016). Learning transitive verbs from single-word verbs in the input by young children acquiring English. *Journal of Child Language*, *43*, 1103-1130.

Serratrice, L., Joseph, K. L., & Conti-Ramsden, G. (2003). The acquisition of past tense in preschool children with specific language impairment and unaffected controls: regular and irregular forms. *Linguistics*, *41-42*, 321-349.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language*, *31*, 61-99.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

To cite this article: Ninio, A. (2017). Pathbreaking verbs for developing recursion in the speech of children acquiring English. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 213-219). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Evaluating the F_0 curve in the speech of individuals with Down syndrome

Marian Oliveira, Vera Pacheco, Maria de Fátima Baia
mdossoliveira@gmail.com, verapacheco@gmail.com, mariadefatimabaia@uesb.edu.br

Universidade Estadual do Sudoeste da Bahia

Abstract. This paper aims at analyzing the F_0 curve of Brazilian Portuguese (BP) high vowels [i] and [u] in stressed and unstressed syllables produced by three young people with Down syndrome (DS). After analyzing F_0 curve of high vowels produced by individuals with DS, we observe that there was no difference in the F_0 curve of all vocalic segments both in unstressed and stressed syllables. Thus, word stress in DS speech does not present identical aspects regarding word stress placement in typical BP. Our explanation for this discrepancy relies on the physiological characteristics of the vocal tract of people with DS as the state of low muscle tone which tends to involve reduced muscle strength (hypotonia).

Keywords: Down syndrome, word stress, Brazilian Portuguese

Introduction

This paper aims at analyzing the F_0 curve of Brazilian Portuguese (BP) high vowels [i] and [u] produced by three young people with Down syndrome (DS). Studies conducted by Oliveira (2011), and Oliveira and Pacheco (2013) observe inter and intraspeaker variability in the formantic patterns of oral vowels produced by children and teenagers with DS. Apart from that, the authors demonstrate that people with DS show difficulty in establishing difference between stressed and unstressed vowels. They relate their findings to aspects of people with DSˈs vocal tract as floppiness (hypotony), small mouth with a tongue that may stick out (macroglossy).

DS people exhibit peculiar physical and physiological characteristics related to their vocal tract: orofacial hypotonia, macroglossia and small jaw. These aspects may make the voice of a DS person sound peculiar and sometimes cause intelligibility problems (Mustacchi & Peres, 2000).

In BP, fundamental frequency (F_0) is an important acoustic parameter to determine voice quality as it is the first frequency produced in the glottis. Additionally, F_0 plays a role in word stress placement in BP together with duration and intensity. The *hypothesis* is that hypotony may have an influence on the vocal folds of people with Down syndrome, which makes the task of word stress placement harder, because the flaccidity alters the vibration of vocal folds and consequently the way word stress placement occurs.

Etiology and language characteristics in DS

Down syndrome is understood as a genetic disorder and the cause of intellectual disability. In general, the risk to have a newborn with Down Syndrome is 1 in 700. In most cases the disorder is caused by trisomy 21, i.e. the presence of three copies of chromosome 21. However, DS can occur by translocation (one part of chromosome is attached to another), and mosaicism when some of the cells in the body are normal and some cells have trisomy 21. In addition, it is common to associate this genetic disorder with maternal age (Hassold & Sherman, 2002).

Martin, Klusek, Estigarribia, and Roberts (2009:1) state that '[…] language and literacy competence may be affected by the cognitive-behavioral phenotype associated with a diagnosis of Down syndrome.' Furthermore, studies conducted by Miller and Leddy (1998) and Miller, Leddy, and Leavitt (1999) report that DS people show variations in their skeletal, muscular, and nervous system. The authors claim that these alterations have an impact on DS people's speech production as 'poorly differentiated midface muscles and/or additional facial muscles may contribute to difficulties with articulation'. Therefore, these variations in the size of oral and pharyngeal cavities may influence how sound travels through the cavities and directly affect the acoustic characteristics of the output. Briefly, reduced size of the oral cavity and a high palatal arch may make the tongue placement for some speech sounds difficult. Similarly, 'a larger tongue in relation to the size of the oral cavity may interfere with articulatory placement.' (Miller & Leddy, 1998:61).

These findings can be found in earlier studies on DS speech, as well. In a historical literature review about DS since the 1950s, Kent and Vorperian (2013) concluded that there are four main areas of investigation in the DS literature: (i) voice, ii) speech production, iii) fluency and prosody, and iv) intelligibility. Studies conducted between 1949 and 1977 on vocal characteristics focus on the evaluation of fundamental frequency (F0), level and quality of voice. These studies are guided by the hypothesis that DS is associated with a characteristic dysphonia, i.e. the voice of a DS person would be characterized as low-voiced, hoarse and rough (Kent & Vorperian, 2013:179). The authors remind us that these studies were carried out with about 600 DS people and have contributed to current research on the linguistic aspects of the speech of children with Down syndrome.

Studies conducted on fluency and prosody indicate the occurrence of disfluency (either stuttering or cluttering), a communication disorder characterized by involuntary speech interruption. The disorder reaches an average of 10 to 40% of the population with DS; in contrast, the occurrence is only at 1% in the population with typical speech (Kent & Vorperian, 2013).

Kent and Vorperian (2013) point out the occurrence of even more significant prosodic disorders and their social consequences. The authors claim that unintelligible speech throughout the lives of people with DS may have negative effect on their social and vocational activities. Although poorly detailed, studies report that disorders of voice, articulation and resonance, which influence fluency and prosody, contribute to the problem.

Nevertheless, Kent and Vorperian (2013:185) state that '[…] disfluency is highly likely to occur in DS but it is by no means a universal characteristic of the syndrome'. Thus, the reported studies present results regarding prosodic limitations in DS, which can be understood as a result of motor and coordinating difficulties as well as speech motor control, but they do not claim these DS speech aspects as a universal bias.

From the review of Kent and Vorperian (2013), we highlight those claims related to phonetic and phonological studies on early language development, which show that DS children may present a delay to start babbling. Due to this articulatory delay, we can observe in their speech abnormality, developmental sounds disorder, supraglottic alterations, distortions in the productions as well as problems in delimitating the contrasts between low/high vowel and front/back vowel.

The emergence and production of consonant oppositions present a delay as well in the development of children with DS. For instance, very productive sounds in English such as / d, t, n, v, l, s, r / (some of which are alveolar, have a high impact on speech intelligibility and are acquired without delay by typically developing children) are not fully acquired by people with DS until 15 and 22 years of age. In addition, other phonetic characteristics are found such as clusters simplification in final and initial word position. The general conclusion is that there is global delay in the phonetic/phonological development of people with DS due to motor limitations and anatomical factors (Kent &Vorperian, 2013).

Studies on Brazilian Portuguese spoken by people with DS, especially the ones conducted by Oliveira (2011), Oliveira and Pacheco (2013), and Oliveira, Pacheco and Pereira-Souza (2017) establish a relationship between the alterations in the vocal tract during DS and in the acoustic quality of vowels and consonants. The authors explain that people with DS have great difficulty establishing the articulatory trajectories of low and high vowels, as well as producing stressed and unstressed syllables. They relate their findings to aspects of the individual's vocal tract, such as hypotonia and macroglossia. Moreover, formant frequencies (F1-F2) are low in low/open vowel and relative higher in high/close vowels. In other words, F1 and F2 values are low in /a/, whereas F1 values are high and F2 values are low in /i, u/.

Lexical stress in BP in typical and atypical speech

According to Cintra (1997), Portuguese has more trochees (63%) than iambs (18%). Child language data suggest an unclear bias, which is not consistent with frequency data in the adult speech. Most of the studies on BP acquisition are naturalistic and they assert that there is an initial iambic bias, while studies on Germanic languages observe a trochaic bias instead (Archibald, 1995; Fikkert, 1994). By analyzing the spontaneous speech of two Brazilian children, Santos (2007) points to a strong initial bias, both in nouns and verbs, in the early words (until 1;6 years old). This result is corroborated by Bonilha (2004), who shows that there is a larger number of iambs in the speech of a Brazilian child until age 1;5. After that age, the child is able to produce the adult prosodic pattern of BP (trochaic words). Following the generative view, Santos (2007) then argues that the difference in results from the Germanic languages is because there is no initial value for the headedness parameter, as claimed by early studies (e.g., Archibald, 1995; Fikkert, 1994). Santos states that children's first words reflect the early setting of the parameter for foot headedness, while extrasyllabicity was not yet acquired. That would be the reason why Germanic languages show an early tendency for trochees, while in BP the early tendency is for iambs.

On the other hand, an experimental study on BP (Rapp, 1994) states that there is a trochaic bias in children's first words, that makes the results differ significantly between naturalistic studies (that point to an iambic bias) and experimental studies (that point to a trochaic bias). An important difference between this study and the naturalistic ones is that it analyzes only one grammatical category, nouns. Therefore, the difference between the results of Rapp on one hand and the those in the studies of Santos and Bonilha raised the question of whether the discrepancy stems from the methodology employed. In order to answer the question, Baia and Santos (2010) conducted an experimental and a naturalistic study to compare the results of each type of study. Interestingly, their findings corroborate what has been reported by both the experimental (Rapp, 1994) and the naturalistic studies (Baia, 2010; Bonilha, 2004; Santos, 2007), that is a trochaic and an iambic bias in early productions, respectively. The authors show that the study on initial word stress and the analyses on early prosodic patterns of BP must take into account all types of data (both experimental and spontaneous and words from different grammatical classes together). The methodological difference found was not due to the kind of elicitation task (experimental vs. naturalistic), but due to the kind of words analyzed. The authors argue that when only nouns are taken into consideration in the analysis, a trochaic bias is found in BP, whereas when both nouns and verbs are considered, an iambic tendency is found. In sum, both types of words, i.e. verbs and nouns, provide full information on Brazilian children's phonological processing during the early stages of acquisition.

As far as we know, this is the first study on word stress in BP atypical speech, in specific in BP spoken by people with DS. Pettinato and Verhoeven (2009), in a study on the lexical stress in Native English-speaking children and adolescents with DS, remind us that children with language disorders tend to show a delay in developing stress patterns because of their complexity. Also, the authors (Pettinato & Verhoeven 2009:3) mention that 'a few studies of language in Down syndrome have included some aspect of prosody (i.e. intonation and rhythmic aspects of language) but none of them have assessed word stress directly'. After analysing experimental data, which consisted of nonsense words repeated by the participants, Pettinato and Verhoeven (2009) observe that children and adolescents with DS have difficulties in both the production and perception of stress patterns, as well as in producing word-initial weak syllables. Thus, we do not expect in our analysis to observe the same aspects of word stress of BP typical speech in the productions of people with DS.

F_0 as one of the acoustic parameters in BP word stress

Massini-Cagliari (1992) states that word stress in BP is characterized by three acoustic parameters: duration, intensity and F_0. Although the author emphasizes the role played by duration on stress placement, Consoni, Arantes, Barbosa, and Ferreira Nett (2006), in a study on the role of each parameter in the perception of word stress of BP, show that it is not easy to state which would be the most important one for the perception of BP word stress. Therefore, we can conclude that both parameters should be taken into consideration in the analysis of BP word stress. However, in this initial study, we start analyzing the speech production of people with DS focusing on one acoustic parameter: F_0. According to Behlau and Pontes (1995), fundamental frequency is physiologically determined by the number of cycles of vocal folds per second and tends to present differences in atypical speech.

Method

In order to investigate the F_0 values of BP high vowels [i] and [u], we analyzed experimental data produced by three female speakers with DS (henceforth S1(M), S2(L), S3(N)), aged 13-18 years, born in Vitoria da Conquista (Bahia, Brazil). In the experiment, the individuals repeated four times the carrier phrase *Digo* _____ *baixinho* with disyllabic nonsense words. After collecting the data in LAPEFF-UESB (phonetics and phonology lab), the extraction of F_0 measurements in pretonic (PT) and tonic (T) syllables was performed in Praat (Boersma & Weenink, 1996). F_0 values of T syllables were contrasted with the values of PT syllables and a Kruskal-Wallis Anova test was performed in BioEstat 5.3. The extraction of F_0 measurements was also performed in on unstressed and stressed syllables in three points of vowels: initial point (P1), medial point (P2) and final point (P3). The F_0 values of unstressed syllables were contrasted with the values of stressed syllables. For this, we used One-Way ANOVA (http://www.socscistatistics.com/tests/anova/default2.Aspx) assuming significant difference when $p \leq 0{,}05$. This research is authorized by UESB Research Ethics Committee.

Results and discussion

Comparison between the mean F_0 of the points of vowels [i] and [u] in unstressed and stressed syllables

As mentioned before, word stress is distinctive and delimiting in BP. Therefore stress production is linguistically significant in Portuguese. The following tables present the average F_0 (Hz) of the participants with DS producing vowel combinations with different consonants: bilabials, postalveolar affricates and velar stops, in nonsense words.

Table 1. [i] average F_0 (Hz) close to bilabials, postalveolar affricates and velar stops in P1, P2 and P3 of unstressed and stressed syllables – SL, SM and SN

Participants	Consonants	POINTS OF VOWEL								
		P1			P2			P3		
		PT syllabes	T syllabes	p-value	PT syllabes	T syllabes	p-value	PT syllabes	T syllabes	p-value
SL	Bilabials	264,6	247,3	0,76 ns	253,1	256,2	0,70 ns	251,3	303,6	0,61 ns
	Postalveolars affricates	275,0	261,3	0,84 ns	260,1	259,6	0,68 ns	241,7	316,0	0,47 ns
	Velars	256,9	271,7	0,87 ns	258,9	271,1	0,87 ns	251,2	325,0	0,53 ns
SM	Bilabials	280,5	270,7	0,92 ns	270,8	266,6	0,97 ns	270,1	221,5	0,94 ns
	Postalveolars affricates	266,7	275,4	0,93 ns	271,7	275,0	0,84 ns	252,5	237,4	0,78 ns
	Velars	271,5	257,9	0,92 ns	298,1	265,8	0,73 ns	251,2	271,7	0,69 ns
SN	Bilabials	236,2	226,1	0,79	253,8	231,0	0,99 ns	252,8	227,3	0,69 ns
	Postalveolars affricates	247,7	225,5	0,66 ns	260,2	232,7	0,95 ns	246,8	230,4	0,85 ns
	Velars	246,3	226,6	0,99 ns	252,7	235,1	0,70 ns	245,3	235,3	0,93 ns

Tables 1 and 2 show that there was no significant difference between the F_0 value of vowels in tonic and pretonic syllables.

Probably, inadequate or insufficient vibration of the vocal folds is reflected in the values of F_0 in people with DS. As word stress is distinctive in BP, the lack of acoustic difference between stressed and unstressed syllables may cause a negative effect on speech intelligibility and, consequently, on communication.

Table 2. [u] average F_0 (Hz) close to bilabials, alveolars and velar stops in P1, P2 and P3 of unstressed and stressed syllables – SL, SM and SN

Participants	Consonants	POINTS OF VOWEL								
		P1			P2			P3		
		PT syllabes	T syllabes	p-value	PT syllabes	T syllabes	p-value	PT syllabes	T syllabes	p-value
SL	Bilabials	251,1	244,1	0,69 ns	248,7	256,7	0,98 ns	242,4	280,7	0,94 ns
	Alveolars	258,4	255,0	0,78 ns	253,2	260,0	0,94 ns	249,0	304,9	0,87 ns
	Velars	247,7	245,2	0,96 ns	249,9	238,2	0,93 ns	272,7	301,2	0,43 ns
SM	Bilabials	281,3	268,3	0,90 ns	281,9	266,5	0,78 ns	257,8	268,9	0,92 ns
	Alveolars	242,2	265,1	0,82 ns	243,0	267,3	0,81 ns	232,0	260,9	0,78 ns
	Velars	267,3	264,7	0,80 ns	270,9	266,4	0,85 ns	250,0	274,3	0,69 ns
SN	Bilabials	236,9	219,8	0,79 ns	240,3	221,4	0,99 ns	238,5	218,9	0,85 ns
	Alveolars	246,0	240,7	0,96 ns	262,2	243,6	0,91 ns	248,3	244,8	0,98 ns
	Velars	244,9	229,2	0,71 ns	236,7	234,6	0,67 ns	226,9	240,2	0,88 ns

The following figures illustrate the average F_0 of vowels in SL:

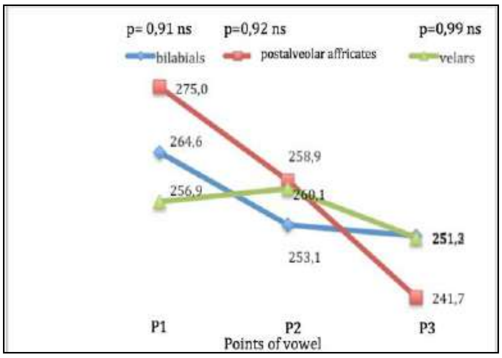


Figure 1. [i] average F_0 (Hz) close to bilabials, postalveolar affricates and velar stops of unstressed syllable – SL

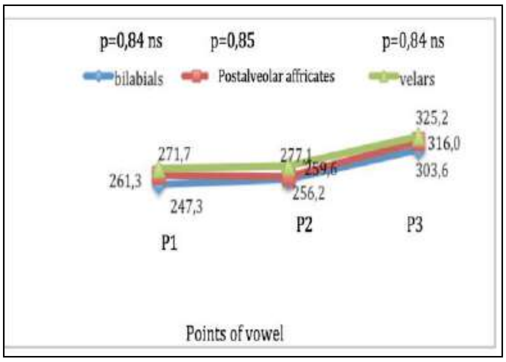


Figure 2. [i] average F_0 (Hz) close to bilabials, postalveolar affricates and velar stops of stressed syllable – SL

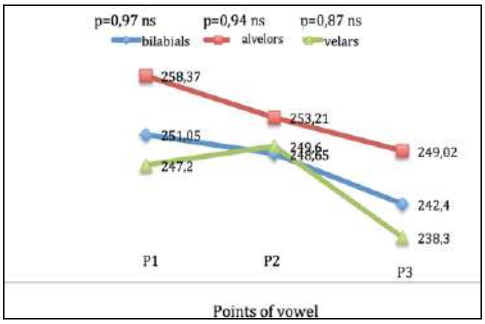


Figure 3. [u] average F_0 (Hz) close to bilabials, postalveolar affricates and velar stops of unstressed syllable – SL

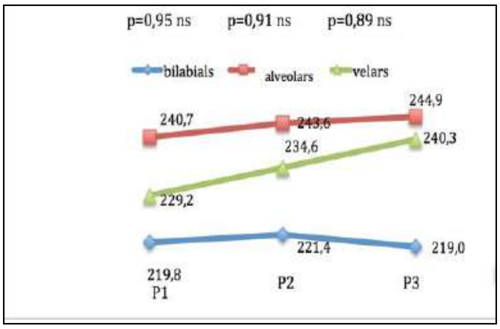


Figure 4. [u] average F_0 (Hz) close to bilabials, postalveolar affricates and velar stops of stressed syllable – SL

After analyzing the F_0 curve of high vowels produced by individuals with DS, we observe that there was no difference in the F_0 curve of any vocalic segments in unstressed syllables and stressed syllables. We claim that people with DS have difficulty articulating the difference between unstressed and stressed syllable due to the hypotonia. For this reason, we claim that hypotonia in individuals with DS may have an influence on their vocal folds, which makes the task of word stress placement more difficult.

In addition, the results point to intra and interspeaker variability which is probably related to variability in the level of muscle flaccidity between the subjects. Flaccidity alters the vibration of vocal folds and consequently the way word stress placement occurs in their speech production. The bias observed by Oliveira (2011) and Oliveira & Pacheco (2013) are confirmed by our results.

Conclusion

Our study continues the investigation of duration and intensity in the BP speech of people with DS, as these acoustic parameters play an important role in word stress placement in BP. Also, a perception test is ongoing. Furthermore, it is important to analyze spontaneous data as there is a debate in the literature on word stress development regarding the influence of experimental or naturalistic method on the results (Baia & Santos, 2010). Therefore, comparisons between experimental and spontaneous data in addition to those between nonsense and real words will provide better understanding of word stress placement in the Brazilian speech of people with DS.

Ackowledgements

This research was funded by *CNPq & MEC* as part of the project *Núcleo Saber Down* – Universidade Estadual do Sudoeste da Bahia.

References

Archibald, J. (1995). The acquisition of stress. In J. Archibald (ed.), *Phonological acquisition and phonological theory*. New Jersey: Lawrence Erlbaum Associates.

Baia, M. F. A. (2010). *O modelo prosódico inicial do português brasileiro: uma questão de metodologia?* São Paulo: FFLCH/USP Produção Acadêmica Premiada.

Baia, M. F. A., & Santos, R. S. (2010). The discrepancy between naturalistic and experimental studies: the case of the initial prosodic template in Brazilian Portuguese. In ed. A. M. Scott (ed.), *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology* (pp. 127-135). Somerville, MA: Cascadilla Proceedings Project.

Behlau, M. S., & Pontes, P. (1995) *Avaliação e tratamento das disfonias.* São Paulo: Lovise.

Boersma, P., & Weenink, D. (2011). *Praat*: *Doing phonetics by computer* [Computer program]. Version 5.1.10, retrieved 22 November 2009 from http://www.praat.org/

Bonilha, G. F. G. (2004). *Aquisição Fonológica do português brasileiro: uma abordagem conexionista da Teoria da Otimalidade*. Ph.D dissertation, Pontífica Universidade Católica do Rio Grande do Sul.

Câmara, Jr. J. M. (1970). *Estrutura da língua portuguesa*. Niterói: Pontes.

Cintra, G. (1997). Distribuição de padrões acentuais no vocábulo em português. *Confluência*, *5*, 83-92.

Consoni, F., Arantes, P., Barbosa, P. A., & Ferreira Netto, W. (2006). Hipóteses acerca do reconhecimento de acento lexical em palavras isoladas. *Estudos Linguísticos*, *35*, 792-799.

Fikkert, P. (1994). *On the acquisition of prosodic structure*. Ph.D Dissertation, Leiden University.

Hassold T., & Sherman S. (2002). The origin and etiology of trisomy 21. In W. I. Cohen, L. Nadel, M. E. Madnick (eds.), *Down syndrome: Visions for the 21st century* (pp. 295-301). New York, NY: Wiley-Liss.

Kent, R. D., & Vorperian, H. K. (2013). Speech impairment in Down syndrome: A review. *Journal of Speech, Language, and Hearing Research*, *56*(1), 178-210.

Martin, G. E., Klusek, J., Estigarribia, B., & Roberts, J. E. (2009). Language characteristics of individuals with Down syndrome. *Language Disorders*, *29*(2), 112-132.

Massini-Cagliari, G. (1992) *Acento e ritmo.* São Paulo: Editora Contexto.

Miller, J. F, Leddy M., & Leavitt L. (eds.). (1999). *Improving the communication of people with Down syndrome*. Baltimore, MA: Brookes.

Miller, J. F., & Leddy, M. (1998). Down syndrome: The impact of speech production on language development. In R. Paul (ed.), *Exploring the speech-language connection* (pp. 163-177). Baltimore, MA: Brookes.

Mustacchi, Z., & Peres, S. (2000) Síndrome de Down. In *Genética baseada em evidencias: síndromes e heranças* (pp. 817-894). São Paulo, CID.

Oliveira, M. (2011) *Sobre a produção vocálica na síndrome de Down: descrição acústica e inferências articulatórias*. PhD Dissertation. Campinas: UNICAMP.

Oliveira, M., & Pacheco, V. (2013). Produção vocálica: análise acústica e síndrome de Down. *Diadorim*, *14*, 99-126.

Oliveira, M., Pacheco, V., & Pereira-Souza, L. (2013). Produção vocálica: análise acústica e síndrome de Down. *Diadorim*, *14*, 99-126.

Pertinatto, M., & Verhoeven, J. (2009). Production and perception of word stress in children and adolescents with Down syndrome. *Down Syndrome Education International*, # 1062823.

Rapp, C. (1994) *A elisão de sílabas fracas nos estágios iniciais da aquisição da fonologia do Português.* Master's Dissertation, Universidade Federal da Bahia.

Santos, R. S. (2007). *A aquisição prosódica do português brasileiro de 1 a 3 anos: padrões de palavra e processos de sândi externo.* São Paulo: Universidade de São Paulo.

Wertzner, H. F., Schreiber, S., & Amaro, L. (2005). Análise da frequência fundamental, jitter, shimmer e intensidade vocal em crianças com transtorno fonológico. *Rev Bras Otorrinolaringol*, *71*(5), 582-88.

To cite this article: Oliveira, M., Pacheco, V., & Baia, M. F. (2017). Evaluating the F_0 curve in the speech of individuals with Down syndrome. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 220-226). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Early lexical development in the acquisition of Turkish: A developmental profile

F. Hülya Özcan
fozcan@anadolu.edu.tr

Anadolu University

Abstract. Crosslinguistic studies on lexical development have shown that both language-general and language-specific characteristics influence the order of acquisition of word classes. In line with language-specific characteristics, previous research, though limited, has shown that in bilingual studies, acquisition of Turkish lexicon draws a different profile. To further delineate this different profile, the aim of this study is to investigate the nature of monolingual Turkish-speaking children's early lexicon with respect to the noun/verb bias phenomenon. For this reason, the vocabulary development of 121 Turkish-speaking monolingual children between 0;8 and 36 months was studied. The data were collected using the Turkish version of CDI-I and CDI-II. Findings reveal that the early lexicons contain more nouns than verbs. Acquisition of verbs accelerates when the size of the lexicon exceeds 50 words. Although more verbs are acquired at this stage, the growing lexicon size is characterized by nouns at later ages. However, the ratio of verbs to nouns becomes smaller with increasing age. This finding has led to a conclusion that a longitudinal study will explain better the order of the acquisition.

Keywords: noun/verb bias, Turkish, first language acquisition

Introduction

Early lexical development is considered to be central in the acquisition process. One important aspect of early lexical acquisition is the order of acquisition. Earlier research has reported that children's productive vocabulary consists of mainly nouns (Goldin-Meadow, Seligman, & Gelman, 1976). The main theoretical argument on this issue came from Gentner (1982), who claimed that early lexical acquisition is noun biased. This noun-biased view argues that children are surrounded by concrete objects and their lexical realizations in everyday life, and this makes nouns perceptually accessible to them. In other words, children's early lexicon comprises of the nouns whose referents are available in the extralinguistic context. Verbs, on the other hand, are characterized by late acquisition. What makes verbs difficult to acquire is the fact that verbs express relational meanings, which makes them conceptually difficult.

Gentner (1982) presented data from six languages, English, German, Japanese, Kaluli, Mandarin Chinese, and Turkish to show that a noun biased early lexicon is universal. Maguire, Hirsty-Pasek, and Golinkoff (2006:367) claim that Gentner's work 'spurred a flurry of activity that spanned 25 years of prolific research from languages that covered the globe', so the existing literature contains studies on English, French, Spanish, Italian, Finnish, Hebrew, Mandarin Chinese, Navajo, Korean, and Turkish.

Crosslinguistic evidence has been provided by studies on different languages, such as Italian, Korean, Hebrew, French, Spanish, Dutch, as well as, English (Bornstein, Cote, Maital, Painter, Park, & Pascual, 2004; Caselli, Bates, Casadio, Fenson, Fenson, Sanderl, & Weir, 1995; Dromi, 1987; Goldfield, 1993; Gillis & Verlinden, 1988; Jackson-Maldonado, Thal, Marchman, Bates, & Guitierrez-Clellen, 1993). Among these, a number of studies confirm Gentner's findings while others challenge the noun bias.

Studies challenging noun dominance can be classified into two groups. One group of studies questions the universal nature of noun dominance. Tardif (1996), Gopnik and Choi (1995), and Gopnik, Choi, and Baumberger (1996) stated that children use more verbs than nouns in Mandarin and Korean.

These contrasting findings imply that language-specific characteristics play an important role in lexical development alongside language-general characteristics.

Language-specific characteristics, which have been initially neglected, refer to the morphological transparency of verbs, the structural saliency of nouns, and phonetic saliency. Morphological transparency facilitates early acquisition of verbs. Structural saliency is concerned with the fact that nouns are prominent in a particular language. For example, in non-pro-drop languages such as English, nouns are highly salient, meaning that nouns cannot be dropped in any case. However, in pro-drop or topic-drop languages, like Spanish, Mandarin, and Turkish, nouns lose some degree of saliency and verbs become salient in certain conditions. Phonetic saliency, as pointed out by Slobin (2001), is concerned with the extra pitch that words in final position have. Therefore, in SOV languages, verbs are in phonetically salient positions.

Crosslinguistic findings suggest that language-specific characteristics are at play in early lexical acquisition and the proportion of nouns and verbs acquired by children changes among noun-friendly or verb-friendly languages. Studies on the acquisition of Turkish are scarce and are mainly in comparative bilingual studies (e.g., Özcan, Altinkamiş, & Gillis, 2016). Kauschke, Lee, and Pae (2007) stated that Turkish children at the ages of 3, 4 and 5 displayed good performance with verbs in naming tasks. Altınkamış (2009) studied early lexical acquisition in Turkish in the spontaneous speech of a number of children and suggested that a wider study with the Turkish version of Communicative Development Inventory (CDI) is needed. In fact, the need for a study in the acquisition of Turkish stems from the linguistic structure of Turkish, since Turkish has different parameter settings, which might lead to a verb biased lexical acquisition.

Turkish is an agglutinating language and is considered to be a verb-friendly language although verbs bear all the information, such as person, tense and aspect.

(1) baş -la -t -tır -abil -ecek -miy -di -n?

‘would you be able to have (someone/something) start?

Despite the agglutinating morphology, morphological transparency leads to correct morpheme ordering.

Turkish is a pro-drop language, which allows null subjects and null objects. That makes verbs salient and decreases the saliency of nouns. On the other hand, there has been a recent on-going discussion claiming that Turkish is not a pro-drop but a topic-drop language. The argument comes from the fact that the use of overt pronominals is pragmatically motivated (Öztürk, 2002). Whether a pro-drop or a topic-drop language, nouns lose a degree of saliency.

Turkish also meets the phonetic saliency. The canonical word order is SOV, which allows the verb to be stressed. The pro-drop nature of Turkish also puts an emphasis on the verb, as well.

Within the framework of the existing literature, the present study has been designed to answer the following research questions.

1. What is the developmental profile in the early lexical development of Turkish?
2. What is the nature of monolingual Turkish-speaking children’s early lexicon with respect to the noun/verb bias phenomenon?

Method

Participants

The participants are 121 monolingual Turkish-speaking children between the ages of 0;8 and 36 months. 45 children (15 girls, 30 boys) belong to the CDI-I stage, and 76 children (30 girls, 46 boys) belong to the CDI-II stage. Since the age span is wide between these months, the sample was further divided into 7 subgroups according to their age.

Table 1. Age groups

Age span	Number of participants	
08-11 months	18	CDI-I
12-14 months	16	CDI-I
15-16 months	11	CDI-I
17-19 months	15	CDI-II
20-24 months	18	CDI-II
25-30 months	18	CDI-II
31-36	27	CDI-II

Data Collection

The data were collected using the standardized Turkish version of the CDI-I (TIGE-I) and CDI-II (TIGE-II) developed by Aksu-Koç, Küntay, Acarlar, Maviş, Sofu, Topbaş, & Turan (2009). CDI-I is used for the children between 08-16 months and CDI-II for the children between 17-36 months. To be consistent with the previous studies on different languages, we included the same noun categories and action words.

The inventory was filled in by parents. Since mothers tend to overestimate the child's performance while fathers, on the other hand, underestimate the child's performance, we asked parents to fill in the inventory together. We also asked them to consult with grandparents or caretakers if the child was with a caretaker during the day. After the inventory was filled in, for a cross-check, we either collected video recordings and/or we had negotiations with the parents. We have both cross-sectional and longitudinal data. For this present study, only cross-sectional data were used for analysis.

Data analysis

The first step in the analysis is to draw a developmental profile of receptive and productive vocabulary. Total receptive and productive vocabulary analysis was conducted in relation to age to reveal the developmental profile.

In order to explain the noun/verb bias, the analysis was conducted both in relation to age and vocabulary size. We analyzed the frequency of verbs and nouns. Mean and median values of nouns and verbs were calculated for each subgroup. In order to reveal any significant differences among the occurrences of nouns and verbs in each age group and among genders, we used paired samples *t*-test, independent samples *t*-test and ANOVA, or their nonparametric correspondents.

Results

Receptive vocabulary

First, we look at the developmental profiles of receptive and expressive vocabulary.

Receptive vocabulary growth starts at 8 months and increases with age. The mean value of receptive words is 42 in the youngest group, it increased to 106 words between 12-14 months and to 137 words between 15-16 months. There is a high degree of variation among the children indicating individual differences. Although until 12 months there are children who do not comprehend any single word, the minimum number of words in the receptive vocabulary increases. As the age increases, the range among the children extends, but this higher range is due to the increasing number of the words in the receptive vocabulary of some of the children. Receptive vocabulary increases steadily starting from 8 months towards 16 months and shows a linear increase as a function of age, which is also statistically significant (X^2=39,5;p=0,01). This statistically significant increase takes place after 11 months ($p<0,05$).

Then, we evaluated receptive vocabulary in relation to gender. For the receptive lexicon, the interaction between lexicon size and gender is not significant (X^2=-0,788; p=0,431). Against the assumption that girls are more talkative than boys, boys have larger vocabulary size until 15 months. Girls catch up with the boys only at 15 months.

Productive vocabulary

Productive vocabulary data come from children whose ages range from 8 months to 36 months. Productive lexicon grows steadily but slowly at first, and then accelerates after 20 months, the age at which a vocabulary spurt takes place. An average child between 8 and 11 months does not produce a word yet as indicated by the median value 0 at this age group. After 14 months, each and every child in our study starts production, although some of them express relatively fewer words. The Turkish CDI-I covers the period between 8 and 16 months and includes 418 items. The highest number of the words expressed by one of the children in our study is 154. That shows that, at 16 months, children in our study can produce 37% of the words in the inventory. That may be because the CDI-I covers the age before the vocabulary spurt starts.

The number of words an average child produces around 17 months, the youngest age covered by CDI-II, is 22. CDI-II includes 711 items. That is, at the start of this period, an average child can only produce 3% of the words included in CDI-II. Starting at the 20-24 month period, productive vocabulary grows. Between 20 to 24 months, the average child can produce 31% of the words; in the period from 25 to 29 months, the average child can produce 49% of the words. This ratio increases up to 92% between 30 to 36 months; this means that the average child can produce 92% of the words given in CID-II. Detailed analysis has shown that the remaining 8% of the expressive lexicon, which is not produced by the oldest age group, is related to either more abstract concepts such as time words, or nouns that are not conceptually available in the extralinguistic environment, such as *coke*, *bacon*, *ham*, *sausages*, *soda*, *ketchup*, since these types of food are regarded as unhealthy by mothers and not consumed at all.

The variation between individual children in the number of the produced words becomes higher between 20 to 24 months and grows in relation to the time. Despite the growing variation, there is linear deveopment in the vocabulary size with increasing age (R^2=0, 6571). The difference in productive vocabulary size is statistically significant across ages (X^2=93,7; df=6; p=0,0001). The first substantial development takes place after 19 months in the 20-to-24 month period, therefore, a significant difference takes place between the 17-to-19-month period and the 20-to-24-month period (p=0,05). Another considerable increase in productive vocabulary size takes place after 30 months in during the 30-to-36 month period. The difference between the 25-to-30 month period and 30-to-36 month period is statistically significant (p=0,0001).

We evaluate gender differences in the productive vocabulary taking the median as the basis of the analysis for two reasons: The first reason is that the number of girls and boys is not equal within the age groups, and the second reason is that the data shows a wide range of scores. Median values show that tha average boy has a smaller productive vocabulary than the average girl, until they complete their first three years of life. At one point, between 12 and 14 months, boys have more words than girls, but in this age range individual variation is at play.

Neither boys nor girls display linear development in the vocabulary size. We also evaluate within-gender differences in relation to age. There is a statistically significant difference among the age groups (F=61,1; df=6; p=0,0001). The girls' production accelerates after 24 months. The difference starts after 20 months with the vocabulary burst and continues until 30 months (p=0,001 and p=0,0001 respectively). There is also a statistically significant difference among the age groups of boys (F=15,2; df=6; p=0,0001). The boys' productive vocabulary accelerates in the period between 17 to 19 months, and this increase continues with increasing age. Although the boys' productive vocabulary size continues to increase with increasing age, the increase becomes significant after 29 months (p=0,002).

Nouns before verbs or verbs before nouns?

Receptive vocabulary

We first look at the noun/verb bias in relation to age differences. Receptive vocabulary at all age groups is characterized by noun-biased vocabulary. Both nouns and verbs increase until 15 months. After 15 months, verbs increase steadily as the age increases. Consequently, children's vocabulary contains more verbs than nouns. There is a statistically significant difference between nouns and verbs in the receptive vocabulary (Z= -4,36; p=0,0001). This developmental difference takes place in the periods between 8 to 11 months and 12 to 14 months (Z=2,32; p=0,02 and Z=-3,18; p=0,001), respectively. With increasing age, the understanding of verbs increases.

At this point, within the receptive vocabulary, we should not ignore another category in CDI-I, expressions. The expressions category includes 27 expressions a child hears in the course of daily routines. They are all formed with the impreative forms of verbs. Understanding these expressions would include understanding of verbs, which are consequently included in the receptive verb repertoire. We reanalyze the verb and noun relationship once again considering the number of expressions understood. These expressions are;

acıktın mı?	'did you get hungry?'	dikkatli ol	'be careful'
ellerini çırp	'clap your hands'	buraya gel	'come here'
öpücük ver	'blow a kiss'	topu at	'throw the ball'
bak	'look'	dur	'stop'

With the inclusion of the understanding of verbs in expressions, the total number of verbs in the receptive vocabulary of the average child increases. There is no statistically significant difference between nouns and verbs in the receptive vocabulary.

Productive vocabulary

As seen in the previous sections, there is individual variation in the receptive and productive vocabulary. This individual variation becomes more effective in the productive vocabulary. Therefore, in order to eliminate the effect of wide range between the total number of words, we choose to analyze noun/verb in productive vocabulary, grouping the children according to vocabulary size.

Productive vocabulary displays a noun-biased profile, as well. Verbs are almost half of the nouns, except for the 101-120 word-stage, when children can produce the same of number of nouns and verbs. What happens at this stage and what affects the production would depend on the detailed analysis of the verbs and nouns acquired at this stage. With growing vocabulary, more nouns are produced than verbs.

Children start with more nouns and fewer verbs at the beginning of the acquisition stage. With increasing vocabulary size, there is a growing number of nouns and verbs. The number of verbs reaches a peak in the 51-100 vocabulary size but decreases again as more words get acquired, in other words, as more nouns are added to their vocabulary. With the growing lexicon, the difference between nouns and verbs diminishes and this finding conforms to the conclusion that linguistic cues become prominent around 24 months. Statistical analysis reveals a statistically significant difference between nouns and verbs until they reach 400-word level (p<0,01).

Discussion and conclusion

The traditional account is that verbs are hard to learn because it is difficult to map a specific event without perceiving the relational concepts the verb encodes. This difficulty is attributed to the difficulty in perceiving the relational concepts underlying the verbs. However, it has also been suggested that even if the relational concepts are given in the verb, children still have difficulty in mapping a verb to an action. Maquire, Hirsy-Pasek, and Golinkoff (2006) claim that mapping from action/mental state to word is considerably more challenging than mapping from object to word. That language makes a difference is a widely accepted fact, however, this difference does not characterize early lexicon, since the typology of a language makes a difference starting from the age of 5 even in verb-friendly languages. That is why Turkish, as a verb-friendly language, seems to have made a difference in some studies but not in all, depending on the age span of the participants.

Another factor that may have an effect on noun/verb learning is variability in word order. Dhillon (2016:58) concludes that 'children acquiring morphology complex topic-drop languages will not initially exhibit a consistent noun or verb bias and will eventually develop a verb bias'. It is, therefore, not a mechanism of learning nouns or verbs but that of learning words. Acquiring verbs or nouns is not a one-time process. Children face linguistic cues and linguistically salient entities but also struggle to learn words on a continuum of concepts. The topic-based nature of Turkish would call for words that may present ostensibly a verb-bias picture in the acquisition of Turkish. These results lead to a need for a longitudinal study in which the emergence of first words and the type of words will be studied in detail.

References

Aksu-Koç, A., Küntay, A., Acarlar, F., Maviş, İ., Sofu, H., Topbaş, S., & Turan, F. (2009). Türkçe'de erken sözcük ve dilbilgisi gelişimini ölçme ve değerlendirme çalışması Türkçe iletişim gelişimi envanterleri: TİGE-1 & TİGE-II. Tübitak'a sunulmuş rapor. Proje no: 107K058.

Altınkamış, F. N. (2009). The developmental paths in Turkish children's early lexical composition. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *18*(2), 415-426.

Bornstein, M. H., Cote, L.R.,Maital,S., Painter,K., Park,S.-Y., & Pascual, L. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, *75*, 115-1139.

Caselli, M.C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L. & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, *10*, 159-199.

Dhillon, R. (2016). Examining the 'noun bias': A structural approach. *Proceedings of the 33rd Annual Penn Linguistics Colloquium*, *16*(1), 51-60.

Dromi, E. (1987). *Early lexical development*. Cambridge, UK: Cambridge University Press.

Field, A. (2000). *Discovering statistics using SPSS*. London, UK: Sage.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (ed.), *Language development: Language, though, and culture* (Vol. 2, pp.301-333). Hillsdale, NJ: Erlbaum.

Gillis, S., & Verlinden, A. (1988). Nouns and verbs in early lexical development: Effects of input frequency? Antwerp: University of Antwerp.

Goldin-Meadow, S., Seligman, M. E. P., & Gelman, R. (1976). Language in the two-year old: Receptive and productive stages. *Cognition*, *4*(2), 189-202.

Gopnik, A., & Choi, S. (1995). Names, relational words, and cognitive development in English and Korean speakers: Nouns are not always learned before verbs. In M. Tomasello & W. E. Merriman (eds.), *Beyond names for things: Young children's acquisition of verbs* (pp.83-90). Hillsdale, NJ: Erlbaum.

Gopnik, A., Choi, S., & Baumberger, T. (1996). Formalism or functionalism? Evidence from the study of language development. In M. Darnell, E. Moravcsik, F. Newmeyer, M. Noonan & K. Wheatley (eds.), *Functionalism and formalism in linguistics: Vol. II. Case Studies* (pp. 317-340). Amsterdam: John Benjamins.

Jackson-Maldonado, D., Thal, D., Marchman, V., Bates, E., & Guitierrez-Clellen, V. (1993). Early lexical development of Spanish-speaking infants and toddlers. *Journal of Child Language*, *20*, 523-549.

Kauschke, C., Lee, H.-W., Pae, S. (2007). Similarities and variation in noun and verb acquisition a crosslinguistic study of children learning German, Korean, and Turkish. L*anguage and Cognitive Processes, 22*, 1-28.

Maguire, M. J., Hirsy-Pasek, K., & Golinkoff, R. M. (2006). A unified theory of word learning:putting verb acquisition in context. In K. Hirsh-Pasek & R. M. Golinkoff (eds.), *Action meets words: How children learn verbs* (pp. 364-391). New York, NY: Oxford University Press.

Özcan, F. M., Altinkamiş, N. F., & Gillis, S. (2016). Early lexical composition of Turkish-Dutch bilinguals: Nouns before verbs or verbs before nouns. In E. Babatsouli & D. Ingram (eds.), Special issue on

‘Monolingual and bilingual speech acquisition across languages’. *Poznan Studies in Contemporary Linguisitcs*, *52*(4), 583-604.

Öztürk, B. (2002). Turkish as a non-pro-drop language. In E. E. Taylan (ed.), *The verb in Turkish* (pp. 239-259). Linguistik Aktuell Vol. 44, Amsterdam: John Benjamins.

Slobin, D. I. (2001). Form function relations: how do children find out what they are? In M. Bowerman & S. C. Levinson (eds.), *Language acquisition and conceptual development* (pp. 406-449). Cambridge, UK: Cambridge University Press.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers’ early vocabularies. *Developmental Psychology*, *32*, 492-504.

To cite this article: Özcan, H. (2017). Early lexical development in the acquisition of Turkish: A developmental profile. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 227-233). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Acquisition of South African English by three- to five-year-old children in Cape Town

Michelle Pascoe, Olebeng Mahura
michelle.pascoe@uct.ac.za, o.mahura@uct.ac.za

University of Cape Town

Abstract. South African English is spoken in South Africa, together with ten other official languages. There is limited information on children's acquisition of South African English and multiple languages in this context. We describe the development of speech in preschool children acquiring South African English in Cape Town. Objectives were to describe (a) the phonological processes and phonemic inventory of participants by language background; (b) the prevalence of children with speech disorders in this sample, and (c) the diagnostic category of participants with speech disorders with reference to Dodd's (2005) framework. We assessed 308 children aged 3;0–5;11 acquiring South African English. They were from different areas in Cape Town covering a variety of socio-economic backgrounds. Children were excluded where languages other than English, Afrikaans or isiXhosa (the three main languages spoken in the region) were spoken. Participants were assessed using subtests of the *Diagnostic Evaluation of Articulation and Phonology* (*DEAP*, Dodd et al., 2002). Phonological processes and inventories of the children are broadly comparable with normative data for English speech acquisition documented in the literature, and as expected a progression was observed with increasing age. First language speakers of Afrikaans and isiXhosa showed some aspects of English phonological development linked to influences of their L1. Approximately 9% of children (n=28) had speech difficulties, with the greatest proportion of these (75%, n=21) phonologically delayed. Limited information about typical speech sound development in South African children presents a challenge for speech and language therapists who need to identify and manage children with speech difficulties. Theoretically the work contributes to knowledge of typical speech development in multilingual contexts.

Keywords: South African English, isiXhosa, Afrikaans, speech sound disorders

Introduction

South Africa is a linguistically and culturally diverse nation, with eleven official languages, as well as many other languages not officially recognized. The official languages include the indigenous Bantu languages: isiZulu, isiXhosa, Sepedi, Setswana, Sesotho, XiTsonga, siSwati, Tshivenda, and isiNdebele, as well as English and Afrikaans. Multilingualism is common and the combination of languages spoken varies between the different provinces. In Cape Town, the main languages spoken are Afrikaans (spoken by 49.7%), English (20.2%) and isiXhosa (24.7%) (Statistics South Africa, 2011). Although only spoken by 9.6% of the country's population, English is considered the *lingua franca*.

There is a lack of information about speech development in South African children, and speech acquisition and the nature and prevalence of difficulties have not been well researched. Although some studies have documented speech sound acquisition in isiXhosa (Maphalala, Pascoe, & Smouse, 2014) and Setswana (Mahura & Pascoe, 2016) much of this work has assumed that children are monolingual speakers. While the phonology of adults who speak South African English, Afrikaans and isiXhosa, is well described together with factors that may influence their speech patterns (Van Rooy, 2008), there is very little information on the acquisition of more than one languages in South African children. There is a pressing clinical need to understand what is typical for our context so that speech and language therapists (SLTs) can (1) identify children with speech sound difficulties and offer support to them and their families, and (2) ensure that they are not 'pathologising' children who are, in fact, typical. This is a challenge for SLTs working in South Africa as the lack of information on speech development means that they often have to rely on data and resources that have been developed for English-speaking children in the United Kingdom (UK) and United States (US) (Pascoe & Norman, 2011).

Speech sound difficulties make up a large proportion of SLT caseloads and are most prevalent in children aged 2 to 5 years (Ruscello, 2008). In the United States the prevalence of speech sound difficulties is estimated at 7.5% in children aged 3 to 11 years (Ruscello, 2008). In South Africa, the prevalence of speech sound difficulties is not known. Pascoe, Le Roux, Mahura, Danvers, de Jager, Esterhuizen, Naidoo, Reynders, Senior, & van der Merwe (2015) estimated that approximately 7% of 3-year-old children in Cape Town present with speech sound difficulties, with a large proportion presenting with phonological delays. The present study builds on the work of Pascoe *et al*. (2015) by investigating speech development and prevalence of speech disorders in a larger sample of children in Cape Town.

Methodology

Aims and objectives

The study aims to describe the development of speech in three- to five- year-old children acquiring South African English in Cape Town. The objectives are to describe:

1. The phonemic inventory and phonological processes of the participants by language background;
2. The prevalence of children with speech disorders in this sample;
3. The diagnostic category of participants with speech disorders with reference to Dodd's (2005) framework.

Participants

The study sample of 308 children aged 3;0-5;11 years and acquiring South African English was assessed over a three year period. Children from a range of different areas in Cape Town covering a variety of socio- economic backgrounds were assessed. Children were excluded where languages other than English, Afrikaans or isiXhosa (the three main languages spoken in the region) were spoken. Table 1 provides an overview of participants.

Table 1. Overview of participants by language background

Age (years; months)	Monolingual (SAE)	Afrikaans Bilingual	isiXhosa Bilingual	Trilingual	Total
3;0-3;11	69	48	25	8	150
4;0-4;11	60	19	3	1	83
5;0-5;11	30	26	12	7	75
	159 (51.6%)	93 (30.2%)	40 (13%)	16 (5.2%)	**308**

Procedure

Children were individually assessed in quiet rooms at pre-schools using the articulation and phonology subtests of the *Diagnostic Evaluation of Articulation and Phonology* (DEAP, Dodd, Hua, Crosbie, Holm, & Ozanne, 2002). In these sub-tests, children are required to name target pictures and produce short strings of connected speech. For the picture-naming task, participants who had difficulty producing the target response were given cues. A hierarchy was followed and included providing a semantic cue first (e.g., 'You use it to tell the time'). This was then followed by forced choice within the relevant category (e.g., 'Is it a watch or a phone?'). The child was then asked to imitate the target word if the first two cues failed to elicit a spontaneous production of the correct response. Real-time transcriptions were later cross-checked, with audio-recordings used for re-transcription purposes. A parental questionnaire was used to obtain information about each child's language background and development.

Dialectal considerations for scoring

It has been recommended that dialectal variation be considered when interpreting results obtained from assessments (Dodd *et al*., 2002). Normative data provided in the DEAP comes from British and Australian children. For our study, we consulted literature regarding adult production of varieties of South African English (e.g., De Klerk & Gough, 2004; Van Rooy, 2008). Features that we expected to find in the speech of the participants included the following, based on data from typical adults, as well as our previous study (Pascoe et al., 2015):

- Alveolar trill /r/ and post vocalic /r/ (Afrikaans and isiXhosa)
- Word final devoicing (e.g.,[dok] for *dog*) (Afrikaans and isiXhosa)
- /θ/ produced as /f/ word finally (e.g., [teef] for *teeth*) (Afrikaans)
- Reduced contrasts between long and short vowels (e.g., *seat*/*sit*) (isiXhosa)
- Fewer central vowels and avoidance of schwa (isiXhosa)
- Vowel raising (e.g., [yis] for *yes*) (English L1 speakers)
- A *kit*/*bit* split, i.e. the words *kit* /kɪt/ and *bit* /bət/ do not rhyme. /ɪ/ is used when it occurs before or after velars, after /h/, before /ʃ/ and word-initially. /ə/ is used elsewhere (English L1 speakers)
- Production of *bath* with a low and fully back /a:/ (English L1 speakers).

Bearing in mind these typical adult productions, we modified the DEAP scoring so that children would not be considered atypical if they showed these features.

Data analysis

Results obtained from each participant were analysed quantitatively and qualitatively. Quantitative analysis included documenting the accuracy with which the children produced segments of words. The accuracy scores calculated are percentage consonants correct (PCC) and percentage vowels correct (PVC). For qualitative analysis we described the participants' phonetic inventories and phonological processes. A 90% criterion was used to determine whether or not a consonant had been acquired. A phonological process was considered present if it occurred five or more times in the sample of individual participants.

Results

Objective 1: To describe the phonemic inventory and phonological processes of the participants by language background (monolingual, bilingual and trilingual)

Table 2 shows the mean PCC and PVC scores, together with consonants that still need to be acquired for each group by language background and age.

As expected, the older participants' phonetic inventories were more complete than younger counterparts. Children aged 5;0-5;11 years have more consonants in their phonetic inventories than the younger groups. This trend can be seen across the different language categories. Only interdental fricatives /θ/ and /ð/ develop beyond 5;11 years. The influence of L1 (Afrikaans and/or isiXhosa) on South African English was evident in many of the participants' speech. Alveolar trill /r/, a dialectal variant of /ɹ/, was favoured by many participants. Trilingual participants used /r/ most frequently, followed by bilingual and a few monolingual participants. Interdental /θ/ was produced as /f/ in the final word position by some children. Word-final devoicing was also observed in many participants' speech. Patterns described in adult vowel production, e.g., vowel raising and avoidance of schwa, were noted in some children.

PCC and PVC scores indicate that accuracy of speech production increases with age in children acquiring one or more speech sound systems. In the 3;0-3;11 year group, PVC scores for isiXhosa bilingual children are lower than those of other language groups.

Table 2. PCC, PVC and consonants still to be acquired by children in each of the study age bands

		3;0-3;11 (n=150)	4;0-4;11 years (n=83)	5;0-5;11 years (n=75)
Monolingual (SAE)				
	PCC	83.82 (*sd* 8.6)	93.98 (*sd* 7.17)	96.05 (*sd* 4.91)
	PVC	95.31 (*sd* 3.84)	99.92 (*sd* 0.42)	99.67 (*sd* 0.71)
	Still to acquire	/ʃ ɵ ð ʒ ʤ ɹ/	/ɵ ð/	/ɵ/
Afrikaans Bilingual				
	PCC	83.94 (*sd* 7.13)	93.63 (*sd* 4.21)	95.08 (*sd* 5.88)
	PVC	95.31 (*sd* 3.1)	99.84 (*sd* 0.69)	99.46 (*sd* 0.86)
	Still to acquire	/z ʧ ɵ ð ʒ ɹ/	/ɵ ð/	/ɵ/
isiXhosa Bilingual				
	PCC	83.5 (*sd* 5.4)	94.33 (*sd* 3.51)	96.83 (*sd* 3.87)
	PVC	89.3 (*sd* 7.1)	100 (*sd* 0)	97.75 (*sd* 2.39)
	Still to acquire	/ɵ ð ʒ ɹ/	/ɵ ð/	/ɵ ð/
Trilingual				
	PCC	87.26 (*sd* 6.4)	87 (*sd* 0)	89.43 (*sd* 2.7)
	PVC	94.66 (*sd* 3.75)	100 (*sd* 0)	96.00 (*sd* 3.79)
	Still to acquire	/ŋ v ʧ ɵ ð ʒ ʤ ɹ/	**	**

*** Data not included due to small number of participants*

A phonological process or speech sound simplification was considered present if it occurred five or more times in a child's speech, or two or more times in the case of weak syllable deletion (based on Dodd *et al*., 2002). In addition, 90% of children in each age and language group had to present with a phonological process for it to be considered typical for children in that category. Guidelines provided in the DEAP were used to identify both typical and atypical phonological processes. Phonological processes found in the speech of children acquiring SAE are summarised in Table 3.

Table 3. Summary of phonological processes used by children in each of the study age bands

	3;0 – 3;11 years (n=150)	4;0 – 4;11 years (n=83)	5;0 – 5;11 years (n=75)
Monolingual SAE	Gliding Cluster reduction Fronting Weak syllable deletion	Gliding	Gliding
Afrikaans Bilingual	Gliding Cluster reduction Fronting Weak syllable deletion Stopping	Gliding Weak syllable deletion	Gliding
isiXhosa Bilingual	Gliding Cluster reduction Weak syllable deletion Stopping	Gliding Cluster reduction	Gliding Cluster Reduction
Trilingual	Gliding Cluster reduction Fronting Weak syllable deletion Stopping	Gliding Cluster reduction	Gliding Cluster reduction

Gliding was the most widely used phonological process, frequently seen in all groups. The isiXhosa bilinguals also used cluster reduction and stopping frequently. Stopping, not present in the speech of monolingual speakers, also seems to be used more by Afrikaans bilingual children, although it was not present in the speech of isiXhosa and Afrikaans language groups after 3;11 years. Cluster reduction persists in the speech of isiXhosa bilingual children, probably beyond 5;11 years.

Objective 2: To describe the prevalence of speech sound difficulties in this study sample

Table 4 describes the prevalence of speech sound difficulties in our sample. Twenty-eight participants were diagnosed with a speech sound difficulty. Overall prevalence of speech disorders in our sample was 9,09%. The monolingual group presented with the highest prevalence of speech sound difficulties (13.2%).

Table 4. Prevalence by age group and language background

Age (years;months)	**Monolingual (SAE)**	**Afrikaans Bilingual**	**isiXhosa Bilingual**	**Trilingual**	**Prevalence**
3;0 – 3;11	10.2% (7/69)	6.25% (3/48)	0% (0/25)	12.5% (1/8)	6.66% (10/150)
4;0 – 4;11	16.66% (10/60)	5.26% (1/19)	0% (0/3)	100% (1/1)	14.5% (12/83)
5;0 – 5;11	13.33% (4/30)	7.69% (2/26)	0% (0/12)	0% (0/7)	8% (6/75)
	13.2% (21/159)	6.45% (6/93)	0% (0/40)	12.5% (2/16)	**9.09% (28/308)**

Objective 3: To describe the diagnostic category of participants with speech disorders with reference to Dodd's (2005) framework

Phonological delay was the main subtype for each age band with most children with difficulties falling into this group: 21 of 28 children (75%). Phonological delay is described by Dodd (2005) as speech that is developing in the typical sequence, but more slowly than would be expected given a child's chronological age. A small number of children were diagnosed as having other types of speech sound difficulties (viz. articulation, and speech disorders of the consistent and inconsistent type, Dodd, 2005).

Discussion

This study provides a tentative summary of the South African English speech development of children aged 3;0-5;11 years. This information is needed by clinicians making decisions about children's speech and trying to discern whether development is proceeding typically. The mono- and bilingual children in our study obtained PCC scores that were not statistically different. This gives clinicians a baseline against which children can be compared when determining the presence of speech difficulties. There is a small set of consonants that children may still be acquiring, and these fit with the literature on English consonantal acquisition in other varieties of English (Shriberg, Austin, Lewis, McSweeny, & Wilson, 1992).

A significant difference was noted in PVC scores of bilingual isiXhosa children, who scored below their monolingual English and bilingual Afrikaans peers. This difference is not indicative of delayed development of accurate production of vowels. Rather, bilingual isiXhosa children seem to be reinterpreting the complex English vowel system as a five-vowel system to match that of their L1 (Van Rooy, 2008). This difference was only observed in the 3;0-3;11 year old group, indicating that children acquiring isiXhosa likely develop an ability to differentiate between the English vowels as exposure to English increases and because of a more developed sound system. This difference is an indication of some of the dialectal variations that SLTs working in South Africa need to consider to avoid over-identification of speech sound difficulties.

As expected, the use of phonological processes decreased with increasing age. The most widely used processes include gliding, cluster reduction, stopping and weak syllable deletion. These processes are considered typical for English speech acquisition (Cohen & Anderson, 2011; Dodd *et al*., 2002). Stopping was common in the bilingual groups, but seems to disappear from the age of 4;0. Weak syllable deletion, common in our sample, may only be eliminated at 4;11 years in bilingual Afrikaans children, later than reported for monolingual English and bilingual isiXhosa children. Cluster reduction in bilingual isiXhosa children seems to persist beyond 5;11 years, possibly as children acquiring isiXhosa prefer to follow the usual open syllable structure of the isiXhosa sound system. IsiXhosa has no clusters in it, apart from loan words from English and Afrikaans, so it may be that these children have more need of the simplification process because of less exposure to these words in their L1.

The prevalence of speech sound difficulties in our sample is estimated at 9.09%. This fits with findings from international studies (Law, Boyle, Harris, Harkness, & Nye, 2000), which indicate a range between 2% and 24.6%. The rate estimated for the current study is slightly higher than the preliminary data reported by Pascoe *et al*. (2015) which only considered one age group. Phonological delay was the most common type of speech difficulty in participants diagnosed with a speech sound difficulty, diagnosed in 75% of children with speech sound difficulties. This is in line with research suggesting that a phonological delay is the most common type of speech difficulty (Broomfield & Dodd, 2004).

The study has several limitations. Where children spoke more than one language we did not assess their speech in both/all their languages. Our trilingual samples were small and results from these groups should be interpreted cautiously. The results were based on a sample of children from one of South Africa's cities, and would need to be generalized to other regions of the country with caution. More studies of such a nature should be conducted to assist in making decisions about adaptations which need to be made when using existing formal assessment tools. Despite its shortcomings, the study provides us with much needed information on the nature of typical development and the occurrence of speech sound difficulties in South African children.

Acknowledgments

Michelle Pascoe and Olebeng Mahura are funded by the National Research Foundation of South Africa. We gratefully acknowledge the support of Jane Le Roux and 4th year BSc (Speech Pathology) students who have contributed to this project: Savannah Senior, Aimee De Jager, Amy van der Merwe, Emily Danvers, Juliette Reynders, Chane Naidoo, Natania Esterhuizen (2015); Michael Geiger, Leigh Davids, Varisha Hutheram, Adilah Allie, Wafeeqah Majiet, Candice O'Neill (2016); Amakhoe Goabas, Nicole Motani, Sheridan Pollet, Herlaine Fynn, Camryn Terblanche, Natasha Wood, Cari Klipp, Aqeelah Adams (2017).

References

Broomfield, J., & Dodd, B. (2004). Children with speech and language disability: Caseload characteristics. *International Journal of Language and Communication Disorders*, *39*(3), 303-324.

Cohen, W., & Anderson, C. (2011). Identification of phonological processes in preschool children's single word productions. *International Journal of Language and Communication Disorders*, *46*(4), 481-488.

De Klerk, V., & Gough, D. (2004). Black South African English. In R. Mesthrie (ed.), *Language in South Africa* (pp. 356-380). Cambridge, UK: Cambridge University Press.

Dodd, B. (2005). *Differential diagnosis and treatment of children with speech disorder (2nd ed.).* London: Whurr Publishers.

Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic Evaluation of Articulation and Phonology.* London, UK: The Psychological Corporation.

Law, J., Boyle, J., Harris , F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of literature. *International Journal of Language and Communication Disorders*, *35*, 165-88.

Mahura, O. O., & Pascoe, M. (2016). The acquisition of Setswana segmental phonology in children aged 3.0-6.0 years: A cross-sectional study. *International Journal of Speech-Language Pathology*, *18*, 533-549.

Maphalala, Z., Pascoe, M. & Smouse, M. (2014). Phonological development of first language isiXhosa-speaking children aged 3;0- 6;0 years: A descriptive cross-sectional study. *Clinical Linguistics and Phonetics*, *28*, 176-194.

Pascoe, M., & Norman, V. (2011). Contextually relevant resources in speech-language therapy and audiology in South Africa--are there any? *The South African Journal of Communication Disorders*, *58*, 2-5.

Pascoe, M., Le Roux, J., Mahura, O., Danvers, E., de Jager, A., Esterhuizen, N., Naidoo, C., Reynders, J., Senior, S., & van der Merwe, A. (2015). *Three-year-old children acquiring South African English in Cape Town*. In E. Babatsouli & D. Ingram (eds.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech* (pp. 277-287). ISBN: 978-618-82351-0-6.http://ismbs.eu/data/ documents/ Proceedings-ISMBS-2015.pdf

Ruscello, D. M. (2008). *Treating articulation and phonology disorder in children*. St. Louis: Elsevier Health Sciences.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The Speech Disorders Classification System (SDCS) Extensions and lifespan reference data. *Journal of Speech Language and Hearing Research*, *40*, 723-740.

Statistics South Africa (2011). Pretoria: pp. 23-25. November 13, 2013 Online: http://www.statssa.gov.za /Census2001/Products/Census_2011_Census_in_brief.pdf

van Rooy, B. (2008). *Black South African English: Phonology*. In R. Mesthrie (ed.), *A handbook of varieties of English*. Berlin: Mouton de Gruyter.

To cite this article: Pascoe, M. & Mahura, O. (2017). Acquisition of South African English by three- to five-year-old children in Cape Town. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 234-240). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Perception of French phonemes by Spanish-speaking children (Venezuela)

Ana Isabel Prato
anaprato10@gmail.com

Université Sorbonne Nouvelle, Paris 3

Abstract. Most foreign language learning theories suggest that learners present difficulties when pronouncing new phonemes of the language learned, mostly if they are similar to the phonemes of their mother tongue. This is usually due to the fact that they do not discriminate the new sounds properly. The purpose of this study is to identify the most common French phonemes Spanish-children usually mix up with their opposite sound, when they have to discriminate them. As a result, there is interference during the learning process, the words are mispronounced, which causes problems when communicating. This research presents the results of a test in which we evaluate perception of opposite phonemes in French in a group of Spanish-speaking children. Initially, we determine and analyse errors (confusion of phonemes) and the difficulties of pronunciation presented. This test of auditory discrimination was designed using the 'Psychopy' program. The participants were 36 elementary school children, aged from 6 to 11 years old. There were 18 male participants and 18 female participants. These children do not speak French, but they are willing to learn it. During the test, the participants had to recognize some French phonemes which are generally confused by Hispanic adults. Most of the youngest children presented difficulties when discriminating the sounds in general. 50% of the participants confused phoneme /u/ with phoneme /y/. More than 40% of the participants mixed phonemes /ɛ/ and /œ/. Confusion of nasal vowels /ɑ̃/ - /ɔ̃/ was up to 55%, and we have to add 8% of the children that did not recognize the difference between these two vowels. In the case of opposite consonants, /s/ - /z/ and /v/ - /b/, only 33% of the participants made errors. More than 50% of the participants confused /g/ with /ʁ/, and 8% did not know if they were different or similar sounds. According to Tomé (1997), adults also get confused with these French phonemes, though the percentages are lower tan in children.

Keywords: perception, phonemes, French, Spanish, interference.

Introduction

Perception is the process by which the sounds of a language are heard, assimilated and understood (Chandy, 2012). The role of perception when learning a foreign language is crucial. Learners recognize language speech through perception. So, at what level will perception of language affect the production of the language learned?

Children create a link between speech perception and production based on perceptual experience and a learned mapping between perception and production (Kuhl, 2007). That is, kids imitate the sounds they hear and are guided by the degree of 'match' between the sounds they produce and those stored in memory.

During the language learning process, there are several factors which influence languge performance, such as the quality of the input offered (native speech), age of exposure to the new language, motivation to learn. This study is interested in examining all these three factors.

Early in life, children discriminate among all the phonetic units of the world's languages By adulthood, this universal phonetic capacity is no longer in place and non-native phonetic discrimination is much more difficult (Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola, and Nelsonet, 2007). So, if something is undeniable, this is that kids start to recognize and learn very small differences between speech sounds early in life. This is the reason why adult learners of a second language cannot achieve the kind of fluency that a young native speaker has. And the present work is going to focus on this transition.

The phonological system of children’s mother tongue is getting settled but, when learning/acquiring another language, the two phonological systems enter in conflict. Most foreign language learning theories suggest that learners present difficulties when pronouncing new phonemes of the language learned, mostly if they are similar to the phonemes of their mother tongue. This is usually due to the fact that they do not discriminate properly the new sounds.

The purpose of this study is to identify the most common French phonemes Spanish-children usually mix up with their opposite sound when they have to discriminate them. As a result, there is interference during the language learning process. On one hand, the words are mispronounced, and on the other hand, this could cause problems when communicating.

For better analysis of this problem, we present the results of an auditory test in which we evaluate perception of opposite phonemes in French in a group of Spanish-speaking children from Venezuela. Initially, we determine and analyze perception errors (confusion of the phonemes). In a further study, we will determine the difficulties of pronunciation presented as a result of this perception confusion. This test of auditory discrimination is interactive and was designed with the ‘Psychopy’ program. The test contains two tasks. In the first one, we presented children two (2) words (minimal pairs). They had to determine whether the words (30 items) sound the same or different. In the second task, we presented children one sound (x) to determine whether that sound was present in the words heard (60 items). During the test, the participants had to recognize French phonemes that are generally confused by Hispanics adults. Theories suggest that children are more capable than adults of discriminating new sounds.

The participants were 36 elementary school children aged from 6 to 11 years old. These kids study in Kidopolis CA, a Language Academy, located in Coro, Venezuela. There were 18 male participants and 18 female participants. We did the test to each one of the students, and a language teacher was helping children with the use of the computer and explaining the procedure of the test. These children do not speak French, but they are willing to learn it in the future. They have studied English at a basic level.

<u>Oppositions studied</u>

1. /u/ - /y/
2. /i/ - /y/
3. /ɛ/ - /œ/
4. /o/ - /œ/
5. /ɑ̃/ - /ɔ̃/
6. /ɛ̃/ - /ɑ̃/
7. /a/ - /ɑ̃/
8. /o/ - /ɔ̃/
9. /ɛ/ - /ɛ̃/
10. /s/ - /z/
11. /v/ - /b/
12. /g/ - /ʁ/

Results show that most of the youngest children presented difficulties when discriminating the sounds, in general, in spite of several language theories affirming that younger kids can discriminate with less difficulty than older ones. The responses by gender did not represent significant differences between girls and boys.

In the auditory test, we juxtapposed phoneme /y/ (not known in the Spanish phonetic system) with the 2 similar vowels in Spanish /u/ and /i/. The results demonstrate that out of 100 of the participants in the study, 50 confused phoneme /u/ with phoneme /y/, like in *loup* - *lu*. In the case of /i/ juxtapposed with /y/, the children had more than 50% correct responses (*lit* - *lu*).

More than 40% of the participants mixed phonemes /ɛ/ and /œ/, like in *sert* - *soeur*. With regard to /o/ and /œ/ (*sort* - *soeur*), good answers were up to 75%. Incorrect answers for nasal vowels /ɑ̃/ - /ɔ̃/ were up to 55%, and we have to add 8% of the children that did not see the difference at all between these two vowels, as in *ment*, *mont*. In the case of nasal vowels /ɛ̃/ - /ɑ̃/, there was not much confusión with more than 80% of correct responses. Children were able to discriminate correctly the words *main* - *ment*. In the case of nasal and oral vowel oppositions, mistakes were rare, as the majority of children could differentiate these sounds. According to Tomé (1997), adults also get confused with these French phonemes, though the percentages are lower in children.

That is because, as confirmed by Kuhl (2007), 'Early in infancy, neural commitment is a "soft" constraint; infants' networks are not fully developed and therefore interference is weak and infants can acquire more than one language.'

In the case of opposite consonants, /s/ - /z/ and /v/ - /b/, only 33% make mistakes. For phonemes /g/ - /ʁ/, more than 50% of the children confused them, and 8% did not know if they were different or similar sounds. In an update of the model, Best (2003) suggests that non-native discrimination declines when phonetic contrasts involve the same articulatory organ (/s z/), as opposed to different articulatory organs (/b t/). As affirmed, the studied population has shown that the confusion of these new similar sounds appears to be an obstacle in discrimination. Consequently, we propose that early, clear and systematic teaching of these new sounds could be presented to children when they start to learn French as a foreign language.

References

Best, C., & McRoberts, G. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*. *46*(2-3), 183-216.

Chandy, M. (2012). Improving perception and language acquisition. Retrieved on 16 Septembre 2017 from: http://www.brighthubeducation.com/teaching-tips-foreign-languages/86262-perception-language-learning-and-methodology/

Esposto, R. (2016). *Guide de conversation: Espagnol Latino-américain.* Lonely Planet, Paris.

Goicochea, N., Tomé, M. (2004). EduFLE.net. Internet Archive: *Fautes typiques des hispanophones lors de l'apprentissage du FLE*.

Kamber, A., &Skupien-Dekens, C. (2010). Phonetic improvement in French as a second language: teaching and evaluation in the multimedia laboratory classroom. *Phonétique, Phonologie et Enseignement des Langues de Spécialité*, *1*, 89-102.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Tobey, N. (2007). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society*, 363, 979-1000.

Tomé, M. (1995). L'enseignement de la prononciation du français pour des débutants espagnols. *Actas del II Coloquio Internacional de Lingüistica francesa* (pp. 31-35). University of Seville, 2-4 November.

Tomé, M. (1997). La perception des phonèmes français par des étudiants espagnols. *Estudios Humanístico*s – *Filología* (Vol. 19, pp. 263-269). University of León.

To cite this article: A. Prato (2017). Perception of French phonemes by Spanish-speaking children (Venezuela). In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 241-243). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

What is a family language? Bilingual and/or monolingual speech in Russian-German immigrant families in Germany

Anna Ritter
info@anna-ritter.eu

University of Regensburg

Abstract. There are different opinions on what is really spoken in immigrant families. Some researchers use the terms *family* or *home language*, meaning only one language or language variety (e.g., Schwartz & Verschik, 2013; Tuominen, 1999). Others emphasise the importance of a combination of languages or dialects used in the same context by the same family members and refer to this combination as *familylect* (e.g., Altenhofen, 1996; Søndergaard, 1991). The present paper aims to explore language use in the speech of Russian-German immigrant families in Germany and focuses on the following questions: How many and what languages are used in the immigrant families? How could the specific way of speaking between family members be characterized? Are there tendencies to monolingual speech or language mixing in the family conversations (cf. Auer, 1999)? The present qualitative study is based on the self-recordings of ten immigrant families, sociolinguistic questionnaires and semi-spontaneous interviews. All the adult informants belong to the large migration and repatriation wave from the territory of the former Soviet Union (Kazakhstan, Tajikistan, Uzbekistan, Ukraine and Russia) to Germany in the 1990s. They had little or no knowledge of German on arrival, mainly using Russian among themselves. The analysis of the data showed that, in all the reviewed conversations, parents and children use both Russian and German. Only a few very short parts of these conversations are in a monolingual mode in Russian. Such tendencies to code-switching, language mixing or loss of Russian were found in different family conversations. On the basis of these findings, the way of speaking in the participating families seems to be a specific family style, emerging out of the reciprocal influence of the family members on each other.

Keywords: bilingual speech, language use in the family, code-switching

Introduction

As a result of numerous waves of migration throughout the world and the rapid growth of mobility in the second half of the 20th century, the number of multilingual individuals rose. This fact made an impact on the communication in families, especially in immigrant families being multilingual. A multilingual family is a community of language practice, where individuals have day-to-day communication with each other and use more than one language with their family members (cf. Lanza, 2009). Recent studies on language use and the language maintenance in families with migrant background show significant differences between immigrant generations (e.g., Akinci & Decool-Mercier, 2010; Backus, 2004; Lambert, 2008; Meng, 2001; Wei, 1994) and even within one immigrant generation (e.g., Barron-Hauwaert, 2011; Kopeliovich, 2013).

However, there are different opinions on what is spoken in multilingual immigrant families. A number of linguists consider only one language and use therefore the terms *family language* (Schwartz & Verschik, 2013), *community* or *immigrant language* (Pauwels, 2005; Schüpbach, 2009), and *home language* (Tuominen, 1999). Others emphasise the importance of a combination of languages or dialects used in the same context by the same family members and refer to this combination as *familylect* (e.g., Altenhofen, 1996:108; Søndergaard, 1991).

These different terms partly overlap and not all of them have a clear-cut definition. Thus, there are still open questions. Based on these deliberations, the present paper aims to explore the language use in multilingual families in the case of immigration and focuses on the following questions: How many and what languages are used in the immigrant families? How could the specific way of speaking between family members be characterized? Are there tendencies to monolingual speech or language mixing in family conversations?

The data of the present study comes from Russian-speaking immigrant families in Germany. Most Russian-speaking immigrants belong to the so-called fourth wave of immigration in the 1990s from the countries of the former Soviet Union (cf. Pfandl, 2000; Zemskaia, 2001). During this period of time about 2.36 million ethnic Russian-Germans (Worbs, Bund, Kohls, & Babka von Gostomski, 2013:28), 226,000 persons of Jewish ancestry (cf. Haug, 2007:8) and 417,000 (own estimation on the basis of public German statistics) of other Russian-speakers, who migrated for different reasons like marriage, work or university studies, arrived to Germany.

The majority of these migrants had little or no knowledge of German on arrival, mainly using Russian among themselves. Among ethnic Russian-Germans, only elder generations (cf. Römhild, 1998) or the generations of great-grandfathers (cf. Meng, 2001; Roll, 2003) had some knowledge of German dialects but not of standard German. Younger generations had at best classes of the standard German as a foreign language at school (ibid.).

After spending 20-25 years in Germany and because of the significant size of the group, Russian-speakers developed their own varied infrastructure, including for example supermarkets with typical food, travel agencies, hairdressers, lawyers and a number of newspapers and journals, published in Russian.

Methodology

Participants and design

The present study is based on the self-recordings of ten immigrant families (33 persons) made during their daily routine at home, sociolinguistic questionnaires filled in by all the participants, and semi-spontaneous interviews conducted with the parents. The recordings of nearly natural family conversations (average length 60 minutes/family), where the observer paradox was reduced (cf. Labov, 1972), were transcribed according to the GAT2 system (Selting, Auer, Barth-Weingarten, Bergmann, Bergmann, Birkner, Couper-Kuhlen, Deppermann, Gilles, Günthner, Hartung *et al*., 2009). The sociolinguistic questionnaires with the basic biographic and linguistic data about every participant and semi-spontaneous interviews complemented the information about language use in the family, family language biography, and language acquisition by the family members.

Table 1. Participant families

family code	first generation	second generation	third generation	intermediate generation	number of participants	countries of origin (parent 1 / parent 2)
FA1	1	1	---	---	2	Russia
FA2	2	1	---	---	3	Kazakhstan
FA3	2	2	---	---	4	Tajikistan/Russia
FA4	2	1	1	---	4	Kazakhstan
FA5	1	1	---	---	2	Germany/Russia
FA6	2	4	---	---	6	Uzbekistan/Kasakhstan
FA7	1	2	---	---	3	Ukraine
FA8	1	2	---	---	3	Kazakhstan
FA9	---	1	---	2	3	Tajikistan/Kazakhstan
F10	1	2	---	---	3	Russia
Total	13	17	1	2	33	---

Eight out of ten participating families were ethnic Russian-Germans, one family had Jewish origin and the last one was a bilingual family, in which the mother came from Russia to Germany through marriage. Apart from the division into families, all participants were divided into immigrant generations in accordance with their age of arrival and their language of education (see Table 1). Almost all parents and grandparents (13) belonged to the first immigrant generation that arrived to

Germany between 18 and 30 years old, after having completed their school and (in many cases) higher education. Two participants (parents from family FA9) arrived to Germany at the age of 14 and 16 years, respectively, and went to school in both countries. Therefore, they cannot be counted as first generation. A number of researchers identify those who migrated at the age 10-18 years as the intermediate generation (Backus, 2004) or generation 1.5 (Remennick, 2017). Nearly all participating children (17) formed part of the second immigrant generation given that all of them had their school education in Germany and only two of them arrived to Germany at the age of 1 and 5 years, respectively. One person belonged to the third immigrant generation.

Analysis

The analysis of the present study is predominantly qualitative. Based on the collected data the analysis method was done out of two perspectives: panorama perspective and close perspective. The panorama perspective is based on the data from questionnaires and interviews and aims to analyse the development of the language situation in a family across several generations. This perspective is closely related to language biography studies (e.g., Franceschini, 2001; Meng, 2001). Four main points of this part of the analysis are biographic data and language acquisition, access to languages via media, tendencies in the family language policy, as well as the language environment of the family.

The close perspective is based on the data from the self-recordings intending to show the current use of languages in the families and to compare the statements of the participants with the recorded examples of their speech. Three main points of the close perspective are conversation practices between family members, cases of functional and/or non-functional code-switching, as well as tendencies in the family conversations in accordance with the typology of code-switching of Auer (1999).

Example

Family FA1 consisted of a nuclear family of mother (MU1) and son (SO1) and as an extended family of the parents of MU1 and her brother with his wife and two sons. This family belonged to ethnic Russian-Germans, but the German language was lost during the Soviet time. MU1 grew up in Russia where she finished a middle school. She studied her L2 English as a foreign language at school. At the age of 18, she came to Germany without any knowledge of German. By the time of the recordings, she was 34.

SO1, born in Germany, was seven years old by the time of the recordings and attended the 1st grade of an elementary school. Although the L1 of MU1 was Russian, she used German only with him since the birth of her son until his fifth birthday. SO1’s grandparents, uncle and aunt, on the contrary, always used Russian with him. Thus, German could be counted as his L1 and Russian as his L2.

Family FA1 lived in a German city with a relatively large Russian-speaking community and infrastructure. Therefore, SO1 was available to visit a nearby centre for Russian language, where he learned the Cyrillic alphabet and read his first texts in Russian. Both MU1 and SO1 watched German TV, as well as cartoons or films in Russian on the Internet.

Concerning the close perspective, this family made two self-recordings, when doing homework, playing and cooking together. The following short extract illustrates typical peculiarities of the communication between MU1 and SO1:

0188 MU1: у тебя такие штаНЫ грязные.
‘you have such dirty pants’

0189 SO1: ich WEIß,
‘I know’

0190 MU1: сними их и одень ЧИСтые.
‘take them off and put on clean (pants)’

0191 SO1: kann ich (.) zu PATrick?
‘can I go to Patrick?’

0192 MU1: нет.
'no'

0193 (--)

0194 SO1: waRUM?
'why?'

0195 (2.7)

0196 MU1: а ЧТО ты будешь у патрика делать?
'and what will you do with Patrick?'

0197 (---)

0198 SO1: ich werde mit ihm SPIElen –
'I will play with him.'

The most important feature of the conversation practices consists of permanent use of both languages. The great majority of MU1's utterances directed to SO1 were in Russian, while he responded to them in German. Smith-Christmas (cf. 2016:65) describes this way of conversation as the dual-lingual or parallel mode paradigm. However, in contrast to the study of Smith-Christmas (ibid.), the utterances of SO1 could not be interpreted as an explicit request to switch to a monolingual German medium. On the contrary, these conversation practices seem to be their tacitly negotiated way of speaking with each other. Moreover, in one of the recordings, MU1 unexpectedly talked to SO1 in German. Surprised by this fact, he himself switched to Russian and asked her if she had addressed him in German.

In the recordings of family FA1, very few cases of code-switching were found, most of them in MU1's utterances. These were functional discourse-related switches, for instance when MU1 changed the topic of the conversation or spoke about typical German realities. On the basis of these rare and clear cases of switching, but considering the fact of the dual-lingual paradigm, there was no tendency either to code-switching or to language mixing according to the typology of Auer (1999) in the conversations of MU1 and SO1.

Results

The results of the analysis in the families were aggregated. Concerning both perspectives, a number of similarities between families, as well as between individuals in accordance with the immigrant generations was found.

The first part of the analysis (panorama perspective) showed that the language biography of every family member may affect the language use of the whole family. For example, parents from four participating families (FA2, FA4, FA9 and F10) belonged to the above-mentioned Russian-speaking infrastructure and used Russian language on a regular basis not only at home, but also at their workplace. This made a positive impact on the use of Russian in their families, especially in the family F10, where the son learned the Cyrillic alphabet on his own wish and started to use Russian in communication with his mother via text messages. Mothers from two families (FA5 and FA6) regularly visited their friends in Russia and Kazakhstan and received guests (parents with children) from these countries. As a result, both stated that their children had obtained friends with whom they had to speak Russian only, because these new friends (from Russia or Kazakhstan, respectively) grew up as Russian-speaking monolinguals.

Other examples were found in families FA3 and FA8, where the children protested against the use of Russian in their presence. In family FA3, two adolescent daughters eagerly learned English and French at school, but denied their Russian roots at home. Thus, their parents tried to use very little Russian, when speaking with them, in order to avoid conflicts. In family FA8, both 8-year-old twin sisters blocked every attempt of their mother to teach them Russian. As a result, she gradually resigned and started to use more German with them.

The use of modern electronic media proved to play an important role for the language use in the families. All parents reported using YouTube and diverse Internet portals to watch films and hear music in Russian. In all but two families (FA3 and FA8), parents used the Internet to make their children acquainted to Russian or Soviet cartoons and children's films. Many of the parents found old friends from their countries of origin via social networks and started to use Russian more intensively, orally (via Skype) and in written form.

The second part of the analysis (close perspective) showed that, in all the reviewed conversations between parents and children, both Russian and German were used actively. All parents used Russian when speaking with their children, for example by explanations, praises, requests and instructions, though to a different extent. As expected, the participating children normally spoke more German than Russian. Nevertheless, in all but two families (FA3 and FA8), children accepted the use of Russian by their parents and had dual-lingual conversations, as in family FA1. Only 5-10% of the utterances of children were in Russian.

In six out of ten families (FA1, FA3, FA5, FA6, FA8 and F10), there were found clear conversation practices in the conversations between different members of the families, meaning who used what languages, when and with whom. For example, the mother in family F10 addressed her 10-year-old son in 95% of analyzed cases in Russian. She stated that she spoke with him Russian at home and outside. On the contrary, the mother in family FA3 addressed her children in 80% of cases in German. In the remaining four families (FA2, FA4, FA7 and FA9), the conversation practices were not clear. Thus, it was not easy to predict what language the children and, especially the parents, would use in their conversations next.

In the conversations of the first immigrant generation, both languages were used on a regular basis. Only few short passages of about 2 to 5 minutes in monolingual Russian mode were found in family FA4. These findings could be explained from the fact that, by the time of recordings, all informants from the first immigrant generation had been living in Germany for more than eight years and were accustomed to using both languages in their daily routine. The informants from the second generation used German in the conversations with each other. However, two short passages (1 minute) in Russian were found in the conversations between the children in family FA7, when they played in the absence of their parents. Both children learned Russian at a centre for Russian language and could read and write in Cyrillic. Moreover, their parents strongly promoted the use of Russian in the family, read Russian books and played together using this language. Thus, the children probably replicated the words of their parents, when playing on their own.

There were found different tendencies in the family conversations according to the typology of Auer (1999). Tendencies to code-switching were found in families FA4 and FA5, as their conversations were characterized by functional participant- and discourse-related code-switching. In families FA2 and FA7, there were found tendencies to language mixing, as their conversations were characterized by a number of cases where the function of switching from one language to the other was not clear. Concerning the conversations of the other six families, there were found no clear tendencies in accordance with the typology of Auer (ibid.). Moreover, the conversations in family FA3 tended towards loss of the Russian language within two immigrant generations and transition to German monolingualism.

Discussion

On the basis of these findings, it was revealed that all the participating immigrant families used more than one language in conversations, and there was no geographic separation between the use of Russian and German. Moreover, every family member had access to both languages at home. On the one hand all families watched German TV and, on the other, all parents and many of the children watched Russian films or cartoons online. However, even the parents did not exclusively use Russian in their conversations with each other. Thus, it is not possible to speak about only one language (Russian or German) as a family, home or community language in the participant families (cf. Pauwels, 2005; Schüpbach, 2009).

Only two out of ten families showed tendencies to language mixing. However, their language use seems to be far from amalgamation to a new fused *lect* (cf. Auer, 1999:321). First, these families were not isolated from the German society and had no possibility to develop their own stable language variety in a larger community of practice. Second, German was definitely the dominant language of the children in these families, as it was their language of education and communication with most of their friends. It seems to be unlikely, that this would change.

Nevertheless, there was a specific way of speaking between the members of the participant families, such as the use of certain Russian words or hybrid words, composed out of two languages. Therefore, this way of speaking could be indicated as a specific family style, emerging out of the reciprocal influence of the family members on each other (cf. Altenhofen, 1996:108; Søndergaard, 1991). The term *family style* was labelled by analogy to the group style found in the conversations of adolescents belonging to the second immigrant generation of Turkish immigrants in Germany (cf. Keim, 2006; 2011).

Conclusion

The analysis of conversations demonstrated diverse ways of the development of language use in immigrant families and even within immigrant generations (cf. Barron-Hauwaert, 2011; Kopeliovich, 2013). However, the general tendency to the loss of Russian by the third immigrant generation despite the vast Russian-speaking infrastructure seems to be the same. This tendency is different to those found in other linguistic communities, where the second and the third generations are particularly inclined to maintain their heritage language, for example in the Turkish community in Lyon/France (cf. Akinci & Decool-Mercier, 2010), in the Italian and Chinese communities in New York/USA (cf. Riehl 2009:69) or the Chinese community in Great Britain (cf. Wei, 1994). As it was mentioned before, the main part of the Russian-speaking community in Germany is formed by ethnic Russian-Germans. Their ancestors moved from Germany to Russia in the 18th and 19th century (cf. Meng, 2001). Thus, this group has originally German roots. This fact explains that not all of them are interested in the preservation of Russian.

At the same time, a multilingual family in a modern world is not isolated from society, notably with the emergence of new media like the Internet, Skype and social networks. All these factors have an influence on the development of the use of both Russian and German by the family members and the development of language situation in the participant families as a whole. In order to prove the tendency to the loss or preservation of Russian and to identify changes in the peculiar family styles, it would be reasonable to conduct a follow-up study with the same participants several years later.

References

Akinci, M. A., & Decool-Mercier N. (2010). Aspects of language acquisition and disorders in Turkish-French bilingual children. In S. Topbaş & M. Yavaş (eds.), *Communication disorders in Turkish in monolingual and multilingual settings* (pp. 314-353). Bristol/Buffalo/Toronto: Multilingual Matters.

Altenhofen, C. V. (1996). *Hunsrückisch in Rio Grande do Sul. Ein Beitrag zur Beschreibung einer deutschbrasilianischen Dialektvarietät im Kontakt mit dem Portugiesischen.* Stuttgart: Franz Steiner.

Auer, P. (1999). From code-switching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism, 3*(4), 309-332.

Backus, A. (2004). Turkish as an immigrant language in Europe. In T. K. Bhatia & W. C. Ritchie (eds.), *Handbook of bilingualism* (pp. 689-724). Malden, MA: Blackwell.

Barron-Hauwaert, S. (2011). *Bilingual siblings: Language use in families.* Bristol/Buffalo/Toronto: Multilingual Matters.

Franceschini, R. (2001). Sprachbiographien randständiger sprecher. In R. Franceschini (ed.), *Biographie und interkulturalität. Diskurs und lebenspraxis* (pp. 111-125). Tübingen: Stauffenburg.

Haug, S. (2007). *Soziodemographische merkmale, berufsstruktur und verwandtschaftsnetzwerke jüdischer Zuwanderer.* Nürnberg: Bundesamt für Migration und Flüchtlinge.

Keim, I. (2006). Der kommunikative soziale stil der ‘türkischen Powergirls’, einer Migrantinnengruppe aus Mannheim. *Deutsche Sprache*, *1-2*(06), 89-105.

Keim, I. (2011). Form und funktion ethnolektaler formen: Türkischstämmige jugendliche im gespräch. In L. M., Eichinger, A. Plewnia & M. Steinle (eds.), *Sprache und integration. Über mehrsprachigkeit und migration* (pp. 157-188). Tübingen: Narr.

Kopeliovich, S. (2013). Happylingual: A family project for enhancing and balancing multilingual development. In M. Schwartz & A. Verschik (eds.), *Successful family language policy. Parents, children and educators in interaction* (pp. 249-275). Dordrecht: Springer.

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Lambert, B. E. (2008). *Family language transmission. Factors, issues, outcomes.* Frankfurt am Main: Peter Lang.

Meng, K. (2001). *Russlanddeutsche sprachbiografien. Untersuchungen zur sprachlichen Integration von aussiedlerfamilien.* Tübingen: Narr.

Pauwels, A. (2005). Maintaining the community language in Australia: Challenges and roles for families. *International Journal of Bilingual Education and Bilingualism*, *8*(2-3), 124-131.

Pfandl, H. (2000). *Erstsprachenverwendung und kulturelle Einstellungen von russischsprachigen Emigrierten mit frühem Ausreisealter in deutschsprachiger Umgebung. Elemente einer Analyse der sprachlich-kulturellen Persönlichkeit.* Graz: Universität Graz.

Remennick, L. (2017). Generation 1.5 of Russian-speaking immigrants in Israel and in Germany: An overview of recent research and a German pilot study. In L. Isurin & C. M. Riehl (eds.), *Integration, identity and language maintenance in young immigrants. Russian Germans or German Russians* (pp. 69-98). Amsterdam & Philadelphia: John Benjamins.

Riehl, C. M. (2009). *Sprachkontaktforschung. Eine einführung.* Tübingen: Narr.

Roll, H. (2003). *Jugendliche aussiedler sprechen über ihren alltag. Rekonstruktionen sprachlichen und kulturellen wissens.* München: Iudicium.

Römhild, R. (1998). *Die macht des ethnischen: Grenzfall russlanddeutsche: Perspektiven einer politischen anthropologie.* Frankfurt am Main: Lang.

Schwartz, M., & Verschik, A. (eds.) (2013). *Successful family language policy.* Dordrecht: Springer.

Selting, M., Auer, P., Barth-Weingarten, D., Bergmann, J. R., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M. *et al*. (2009). Gesprächsanalytisches transkriptionssystem 2 (GAT2). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, *10*, 353-402.

Smith-Christmas, C. (2016). *Family language policy: Maintaining an endangered language in the home.* Basingstoke: Palgrave MacMillan.

Søndergaard, B. (1991). Switching between seven codes within one family - a linguistic resource. *Journal of Multilingual and Multicultural Development*, *12*, 85-92.

Schüpbach, D. (2009). Language transmission revisited: Family type, linguistic environment and language attitudes. *International Journal of Bilingual Education and Bilingualism*, *12*(1), 15-30.

Tuominen, A. (1999). Who decides the home language? A look at multilingual families. *International Journal of the Sociology of Language*, *140*, 59-76.

Worbs, S., Bund, E., Kohls, M., & C. Babka von Gostomski (2013). *(Spät-)aussiedler in Deutschland. Eine analyse aktueller daten und forschungsergebnisse.* Nürnberg: Bundesamt für Migration und Flüchtlinge.

Wei, L. (1994). *Three generations, two languages, one family. Language choice and language shift in a Chinese community in Britain.* Clevendon, UK: Multilingual Matters.

Zemskaia, E. A. (2001). *Jazyk russkogo zarubez’ja: obshie prozessy i rechevye* portrety [The Language of the Russian Diaspora: General Processes and Speech Portraits]. Moscow & Vienna: Studia Philologica.

To cite this article: Ritter, A. (2017). What is a family language? Bilingual and/or monolingual speech in Russian-German immigrant families in Germany. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 244-250). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

L2 acquisition and L1 attrition in Bosnian-German late bilinguals - Phonetic observations at the segmental level

Carolin Schmid, Sylvia Moosmüller
carolin.schmid@oeaw.ac.at, sylvia.moosmueller@oeaw.ac.at

Austrian Academy of Sciences

Abstract. The paper investigates the pronunciation of L1 Bosnian and L2 German laterals by Bosnian speakers living in Vienna. Thereby, we make a contribution to the research on linguistic processes of language contact. Two main processes are analysed: Firstly, we investigate whether the speakers maintain the L1 Bosnian phonemic contrast between /l/ and /lj/ in a German speaking environment with only one lateral phoneme. Secondly, we investigate if the similar lateral phoneme category /l/ in L1 and L2 will undergo phonetic changes in one or both languages, since the velarized variant is more widespread in Bosnian, and the alveolar lateral is more prestigious in German. In the present study, spontaneous as well as read speech of 5 women and 3 men is analysed. The potentially influencing parameters 'position of the lateral within the word', 'syllable stress', 'phonemic context', 'recording task', and 'speaker sex' are taken into account. Results show that the two lateral L1 Bosnian phonemes, /l/ and /lj/, differ significantly. Concerning the similar phoneme category /l/, there is also a significant difference between a more velarized lateral in L1 Bosnian and a more alveolar lateral in L2 German. Women and men do not perform differently, and there is no significant difference concerning the speaking task. In a following study, we will analyse control groups of L1 Bosnian and L1 German speakers. Only then we will be able to conclude whether L2 pronunciation is targetlike or not, and whether L1 pronunciation undergoes attrition processes or not.

Keywords: L2 acquisition, L1 attrition, laterals, acoustic phonetics, Bosnian, German

Introduction

In Europe, language contact and bilingualism is becoming more and more widespread, through foreign language teaching at school and, especially, through migration. Researchers keep asking numerous questions about language contact. How do the languages influence each other? What are the reasons for the way and strength of the influence. Can we draw conclusions about identity and integration? Practical questions are also being investigated. Is it cognitively, economically, or socially an advantage or a disadvantage to speak more than one language? Are there differences between different language combinations? What are the linguistic levels which request special attention when learning a second language? Especially the phonetic level is interesting, because it is less conscious to the speaker and more anchored in language use than other linguistic levels. This is because relatively few distinctive sounds exist in one language, and they remain relatively stable over time and over different speaking styles, in contrast to the vocabulary or syntax of a language (Schmid, 2011). Generally, it has been shown that languages in contact spoken by a bilingual/multilingual speaker mutually influence each other at the diverse linguistic levels (see the concept of Crosslinguistic Influence (CLI) by Sharwood-Smith, 1983). In the following, we will shortly summarize some of the relevant phonetic transfer processes described in the literature. Concerning L2 acquisition, merging is often described (at first by Flege, 1987). In this process, an intermediate version of a sound in L1 and a sound in L2 is pronounced in both languages, when the two original sounds share similar articulatory settings. Another process in L2 acquisition is substitution. This happens when a sound in L2 is articulatorily new. An example of this process is: L1 Dutch learners of English who often substitute the English interdental fricative phoneme <th> by an alveolar fricative or a labiodental fricative (Wester, Gilbers, & Lowie, 2007). Concerning the influence of an L2 on the L1, sound shifting, the replacement of a sound in L1 by a variant of this sound in L2, is often observed. One example is the reduction of /o/ in unstressed positions to /a/ in the German pronunciation of L1 German speakers living in Russia (Rosenberg, 2013). Recently, Leeuw, Tusha, and Schmid (2017) found indications for an even stronger sound shifting process. Some of the L1 Albanian speakers in their study lost the lateral phoneme contrast in L1 through contact with L2 English in which the laterals are not distinctive. Mainly time-related factors are mentioned as important parameters for acquisition and attrition, as age of acquisition, age of emigration, length of residence, or amount and quality of contact (for an overview see Schmid, Lahmann, & Steinkrauss, 2014). Additionally, identification (Riehl, 2004), orthography (see Bassetti, Escudero, & Hayes-Harb, 2015, for a hindering effect of orthography concerning a targetlike L2 pronunciation), or gender ('female L2 learners may be more concerned about pronunciation accuracy than their male counterparts', Moyer, 2016), could be important parameters.

The present study is part of the ongoing first author's PhD thesis on sociophonetic aspects of lateral variants produced by Bosnian speakers living in Vienna. Here, we will exclusively focus on the acoustic aspects of L1 Bosnian and L2 German laterals produced by a subset of Bosnian speakers. The Bosnian speakers are mainly exposed to two languages in Vienna: L1 Bosnian and L2 German, more specifically, Standard Austrian German (SAG). Additionally, they may be exposed to the Viennese dialect (Vd), which is spoken by speakers of the lower social classes of Vienna. In Bosnian, two lateral phonemes are distinguished: a palatalized and a velarized lateral phoneme. The velarized lateral shows an alveolar allophone (Petrovic & Grubisic, 2010). The two phoneme categories are also distinguished orthographically: the palatalized lateral is transcribed <lj>, whereas the velarized lateral is transcribed <l>. SAG, on the other hand, features only one lateral phoneme, which is the alveolar lateral. The Vd shares the lateral phoneme inventory with SAG, but in addition, the Vd alveolar lateral phoneme has a velarized variant, which can occur at word initial and word final positions and between back vowels. This velarization is a salient feature, negatively evaluated by Austrian listeners and thus avoided, especially by female speakers (Moosmüller, Schmid, & Kasess, 2015). In the present study, two central questions are asked. Firstly, we will investigate what happens in L1 and L2 when there is one similar phoneme category in the two languages which, have however, phonetically different variants. Secondly, we will analyse what happens in L1 when there is a phoneme contrast in L1, but not in L2.

According to the literature, the best way to describe the lateral quality acoustically is measuring F2 (Thomas, 2011) and F2-F1 to normalize individual vocal tract differences (see de Leeuw *et al*., 2017). All laterals are built with a closure at the mid-sagittal line of the vocal tract, and the airstream passes laterally from the sides of the tongue. In the case of the alveolar lateral, this is mostly a closure between the blade of the tongue and the alveolar ridge. Alveolar laterals have an F2 above 1500 Hz. The darker quality of the velarized laterals is caused by a longer cavity behind the closure in the vocal tract, produced by a constriction in the velar region. The longer the cavity behind the closure, the lower the F2. In velarized laterals, F2 is generally below 1200 Hz (F2 values according to Recasens, 2012) and F1 is increased as compared to alveolar laterals. The articulation of the palatalized lateral is characterized by larger closure at the palatum, which implies higher F2 values (above 2000 Hz), whereas the F1 stays rather low.

With regard to the literature reviewed in the introduction, we formulate the following hypotheses:

1. Bosnian speakers will realize more velarized laterals in Bosnian than in SAG, but the differences will not be significant.
2. There will be more velarized laterals in read speech than in spontaneous speech.
3. Women will perform better than men concerning a targetlike L2 pronunciation.
4. The contrast of /l/ and /lj/ in L1 Bosnian will be reduced, in favor of the /l/ pronunciation.

Method

Speakers

Five female and three male speakers were recorded. Speaker groups as homogeneous as possible were created: All speakers were born and raised in the region of today's Bosnia and emigrated to Vienna during the war in Bosnia, at the age of 20 to 35 years, thus after puberty and having acquired a relatively stable knowledge of their L1 Bosnian. Vienna is still their place of residence at present. The speakers had a similar educational background: they had all started university studies in Bosnia before their emigration to Austria.

Recordings

Read speech as well as spontaneous speech was elicited in both L1 Bosnian and L2 German, although the tasks differed for each language. Read speech was recorded in order to control the lateral context. The sentences were constructed in such a manner that they were as natural as possible, to avoid monotonous or stagnant speech. In the German sentences, the laterals in the target words were balanced for the position of the lateral within the word (initially, medially, finally), for the vowel context (between front vowels, between back vowels, after a front vowel, before a back vowel, and after a back vowel, before a front vowel), and also for syllable stress (see Table 1).

Table 1. Examples for target words and sentences in the German reading task

words	Sentences	English translation
leben	S**ie le**ben nur einmal.	They only live once.
lachen	D**a la**chen ja alle.	Everyone laughs.
Volumen	Das V**olu**men ist groß.	The volume is large.
belebter	Das ist ein b**ele**bter Ort.	This is a busy site.

In total, 110 German sentences were read twice by each speaker. Because Bosnian has two lateral phonemes, a different approach was chosen to select the target words and to construct the sentences. It was likewise attempted to balance the laterals for vowel context and position within the word, but the main focus was on finding minimal pairs with the two lateral phonemes. The target words were embedded in carrier phrases, which were in turn embedded in a short dialog, for the sake of comprehension. This was necessary because it was not always possible to find high-frequency words. Because of the structure of the carrier phrases, word initial laterals were preceded by <a> and word final laterals were followed by <s> (see Table 2). The speakers had to read 36 sentences twice. The recording of spontaneous speech allowed us to analyze the pronunciation of laterals in a less controlled, but more natural setting, and thus to verify or negate the results obtained in the controlled reading task. Spontaneous speech in L2 German was recorded by means of a semi-structured biographical interview with the first author of this paper (an L1 German speaker). The interview included questions about social networks and language use and attitudes. The information obtained in the biographical interviews will be used in further studies for qualitative analyses of the pronunciation performance of the Bosnian speakers. L1 Bosnian spontaneous speech was elicited through three picture stories the speakers had to retell. This method was chosen, because it was not possible for the interviewer to conduct a native language interview in Bosnian.

Table 2. Examples for target words and sentences in the Bosnian reading task

Minimal pairs	Carrier sentences	English translation
ljudi	Da, ljudi sam rekao	Yes, I said people.
ljudi	Da, ludi sam rekao	Yes, I said mad.

Data preparation

The transcriptions of the recordings, as well as the segmentation of the laterals and the words in which the laterals occurred were performed manually in STx (Balazs, Noll, Deutsch, & Laback, 2000). The lateral segment borders were verified through intensity variation, formant transitions, and changes in the waveform. Formant transitions were not excluded from the segment. Formant frequencies (F1-F3) were measured using LPC, then verified and, if necessary, manually corrected. The mean values for each formant were then extracted. Additionally, segment information concerning the phoneme context, the position of the lateral within the word, speaker sex, recording task, and language was annotated and extracted.

Statistical analyses

In a first step, a mixed effects model was carried out in R (R Core Team, 2016) in order to analyse and compare the pronunciation of graphemic <l> in L1 Bosnian and L2 German. Therefore, the difference in Hertz between F1 and F2 (F2-F1) was modelled as a function of the fixed effects ‘speaker sex’, ‘task’, ‘language’, ‘position within the word’, and ‘phoneme context’, and of the random effects ‘speaker’ and ‘word’. The model was built with the lme4 toolbox in R (R Core Team, 2017; Bates, Maechler, Bolker, & Walker, 2015). Post hoc tests were performed using the least-squares means function (lsmeans, see Lenth, 2016) from the lmerTest package (Kuznetsova, Brockhoff, & Christensesn,, 2016). P-values were Bonferroni-adjusted. In a next step, the model was modified in order to measure the relationship between F2-F1 and the two Bosnian lateral phonemes /l/ and /lj/. As this analyse is only relevant for L1 Bosnian (with regard to an eventual reduction of the phonemic contrast), the L2 German productions were excluded from the data for this model.

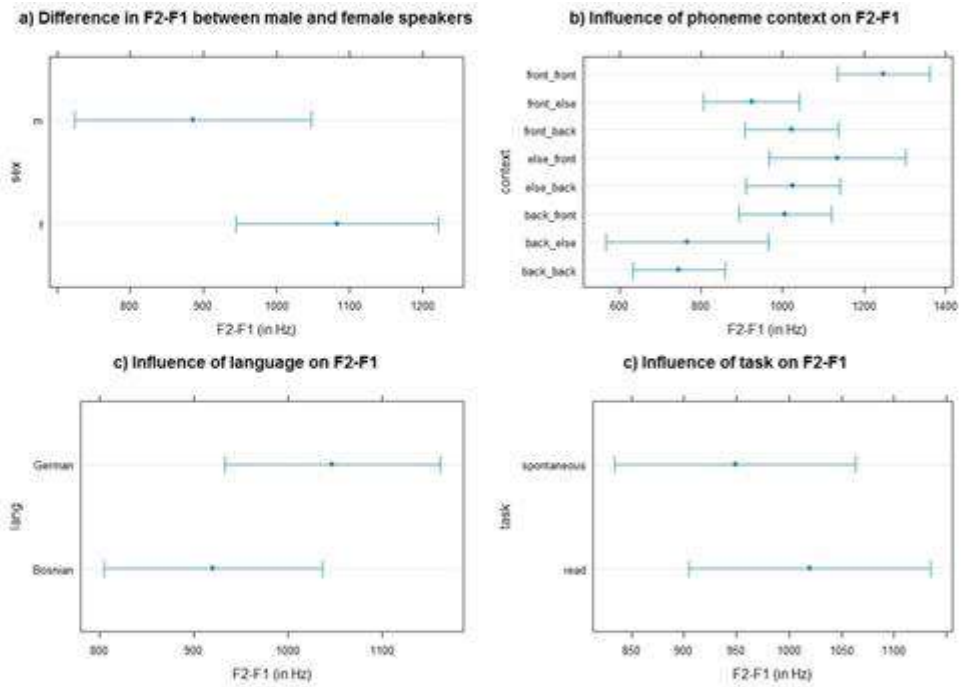


Figure 1. Parameters having a significant effect on F2-F1 of /l/

Results

The analysis of the first model describing the pronunciation of <l> in L1 Bosnian and L2 German showed main effects for speaker sex, phoneme context, language, and task (see Figure 1). The lateral position within the word turned out not to be significant. Women tend to realize laterals with a higher mean F2-F1 than men, as shown in Figure 1a ($F(1,6)=4.61$, $p=0.075$). Figure 1b illustrates that the phoneme context also affects F2-F1 ($F(7, 1255.36)=95.91$, $p<0.001$). A posthoc test reveals that in back_back context, laterals are produced with significantly lower F2-F1 values ($p<0.001$) than in all other contexts, except for the comparison between back_back and back_else context, which doesn’t show a statistical difference ($p=1$). The pronunciation of laterals in front_front contexts is differing most from the pronunciation in back_back contexts and is also significantly different from all other phoneme contexts ($p<0.001$), except from the else_front context ($p=1$). In Figure 1c, the influence of language on the F2-F1 values of laterals is shown. The speakers realize laterals in Bosnian with significantly closer F1 and F2 values than in German (with a mean difference of 126.53 Hz, $F(1,539.87)=49.2$, $p<0.001$). Even though the difference is significant, this result does not show a categorical contrast, as shown by the results of a density distribution we carried out additionally (see Figure 2). Another variable significantly influencing the realization of laterals is the recording task. As shown in Figure 1d, speakers produce laterals in both languages with closer F1 and F2 in spontaneous speech, thus darker than in read speech ($F(1,617.79)=25.86$, $p<0.0001$). Another finding of the model is that there is a significant interaction between task and sex ($F(1, 2834.16)=4.21$, $p=0.04$). But a closer look by means of the posthoc test shows that the difference between female and male speakers is not significant, and that the result is only slightly more significant for male than for female speakers ($p<0.0001$ vs $p=0.0023$).

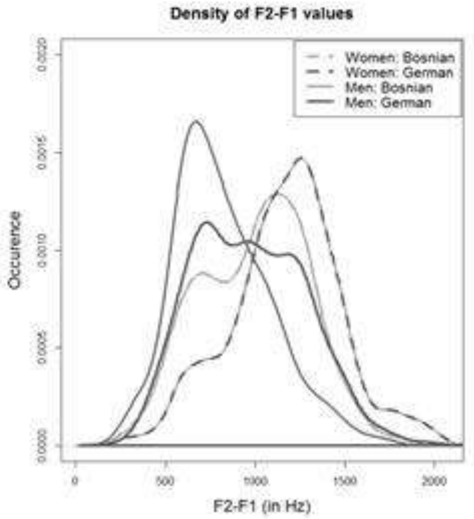


Figure 2. Distribution of F2-F1 values, split into values for women, men, Bosnian and German language

In a next step, the realization of the two lateral phonemes /l/ and /lj/ in Bosnian speech was investigated. As shown in Figure 3, F2 is the most differentiating parameter for the distinction of the two phonemes in the speech of both female and male speakers (F2 is higher in /lj/ than in /l/), but also F1 and F3 differ (F1 is lower and F3 is higher in /lj/ than in /l/). A descriptive analysis carried out initially showed that the difference in Hertz between F1 and F2, as also used for the description of the /l/ laterals in the preceding model, is significantly different for /l/ and /lj/ with a mean of 822.17 Hz for male speakers and a mean of 997 Hz for female speakers in the pronunciation of /l/, and a mean of 1613.43 Hz for male speakers and a mean of 1755.86 Hz for female speakers in the pronunciation of /lj/.

A second mixed effects model was fitted in order to analyse the phoneme contrast of /l/ and /lj/, with the random factors speaker and word, and the fixed factors speaker sex, task, and context. The fixed factor language was irrelevant, because only laterals in L1 Bosnian were considered. Instead, the lateral phoneme was introduced as a fixed factor. All of this factors were again significant, as in the previous model, and also the lateral phoneme turned out to be a significant predictor for the F2-F1 values of the lateral (the lateral phonemes differ with $F(1,147.47)=495.11$, $p<0.001$).

Discussion

In this paper, the pronunciation of laterals of bilingual Bosnian-German speakers living in Vienna was investigated. We hypothesized that the difference between the lateral sounds in the similar lateral category /l/ in L1 Bosnian and L2 German would not be significant, due to either substitution or merging. In our data, this hypothesis could not be confirmed. Speakers still maintain a significant difference between the lateral sounds in their L1 and L2. Laterals in Bosnian speech have closer F2-

F1 values, are thus more velarized than their German counterparts. Nevertheless, the formant values for L1 Bosnian and L2 German laterals are partly overlapping. This is mainly due to the influence of the phoneme context. Next to front vowels, laterals have significantly higher F2-F1 distances than laterals next to back vowels. This result is consistent with what is found in previous studies on the pronunciation of laterals (see for example Moosmüller *et al*., 2015, Recasens, 2012), and is thus not surprising. It will be necessary to conduct further studies with control groups of monolingual Bosnian and monolingual SAG speakers to conclude whether this significant difference found in the present study implies that the speakers have a targetlike pronunciation in both languages, or whether there is nevertheless an influence in one or both directions. In our second hypothesis we stated that orthography will have a hindering effect on a targetlike pronunciation in L2, so that the pronunciation in read L2 speech would be closer to the pronunciation in L1 than the pronunciation in spontaneous L2 speech would be. Our results do not show a difference between the speaking tasks in the expected, but rather in the other: in read speech, laterals are produced with higher F2-F1 differences than in spontaneous speech in both languages. This result is surprising, especially in L2 SAG, where the spontaneous speech was elicited through an interview with a German native speaker who didn't pronounce velarized laterals (see the literature on accommodation effects, e.g., Giles, 2016). Our third hypothesis was not confirmed either, insofar as we found that there are indeed the expected formant frequency differences between female and male speakers (due to different sizes in vocal tract length), but that they did not differ concerning the performance in L1 or L2 or in the different recording tasks. In order to confirm this result, we will record additional female and male speakers in following studies and also qualitatively analyze language attitudes and look deeper into inter-individual differences and their possible reasons. In the fourth hypothesis, we assumed that the phonemic contrast between /l/ and /lj/ in L1 Bosnian will be reduced. At this time, we cannot definitely say

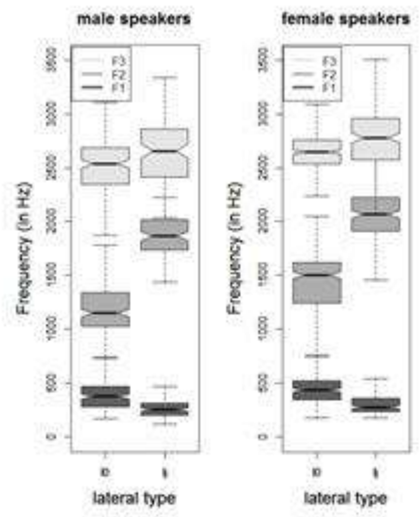


Figure 3. F1, F2, and F3 values for <l> and <lj>, for male and female speakers

whether there is a tendency in this direction. Results only show that the overall contrast is still significant. Here too, the analysis of recordings of a control group of monolingual L1 Bosnian speakers will shed light on possible changes. Overall, the results of the present study show that the influence of one contact language on the other is not so strong as to be directly measurable. The similar lateral phoneme category /l/ in L1 Bosnian and L2 German is still pronounced significantly differently in both languages and the phoneme contrast of /l/ and /lj/ in L1 Bosnian speech is still significant. These results hold for both speaker sex and also for the two recording tasks, and are independent of phoneme context or position of the lateral within the word. In a following study, we will analyse recordings of control groups of monolingual Bosnian speakers and monolingual SAG speakers, in order to measure the influence of the two contact languages more thoroughly. An influence on the phonetic level can still be expected according to the hypothesis of CLI, even though the pronunciation of the laterals in both languages is still significantly different. Additionally, more detailed analyses of interspeaker variation is needed, since it has been shown that a lot of parameters potentially influencing the speech performance in the two contact languages exist (Schmid *et al*., 2014), and thus, speakers can perform very differently (Leeuw *et al*., 2017). Therefore, a qualitative analysis of the data collected by means of the semi-structured biographical interview is needed for following studies.

References

Balazs, P., Noll, A., Deutsch, W., & Laback, B. (2000). Concept of the integrated signal analysis software system STx. *Jahrestagung der Österreichischen Physikalischen Gesellschaft 2000*.

Bassetti, B., Escudero, P., & Hayes-Harb, R. (2015). Second language phonology at the interface between acoustic and orthographic input. *Applied Psycholinguistics*, *36*(01), 1-6.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Flege, J. E. (1987). The production of <new> and <similar> phones in a foreign language. Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*, 47-65.

Giles, H. (ed.). (2016). *Communication accommodation theory: Negotiating personal relationships and social identities across contexts*. Cambridge, UK: Cambridge University Press.

Kuznetsova, A., Brockhoff, P. B., & Christensesn, R. H. B. (2016). *lmerTest: Tests in linear mixed effects models: R package version 2.0-32*.

Leeuw, E. de, Tusha, A., & Schmid, M.S. (2017). Individual phonological attrition in Albanian–English late bilinguals. *Bilingualism: Language and Cognition*, *3*, 1-18.

Lenth, R.V. (69). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, *1*, 10-33.

Moosmüller, S., Schmid, C., & Kasess, C. (2015). Alveolar and Velarized Laterals in Albanian and in the Viennese Dialect. *Language and Speech, 58*(4), 1–28.

Moyer, A. (2016). The puzzle of gender effects in L2 phonology. *Journal of Second Language Pronunciation, 2*(1), 8-28.

Petrovic, D., & Grubisic, S. (2010). *Fonologija srpskoga jezika*. Beograd: SANU.

R Core Team. (2016). *R: A language and environment for statistical computing*.

Recasens, D. (2012). A cross-language acoustic study of initial and final allophones of /l/. *Speech Communication*, *54*, 368-383.

Riehl, C. M. (2004). *Sprachkontaktforschung: Eine Einführung* (3rd ed.). Tübingen: Narr.

Rosenberg, P. (2013).Vergleichende Sprachinselforschung: Sprachwandel in deutschen Sprachinseln in Russland und Brasilien (273-323), *Linguistic Online*, *13*(103).

Schmid, M. S. (2011). *Key topics in sociolinguistics*. *Language attrition*. New York, NY: Cambridge University Press.

Schmid, M. S., Lahmann, C., & Steinkrauss, R. (2014). Sprachverlust im Kontext von Migration und Asyl. In D. Bischoff, C. Gabriel & E. Kilchmann (eds.), *Edition text + kritik: Vol. 32. Sprache(n) im Exil* (pp. 121-131). München: Richard Boorberg Verlag.

Sharwood-Smith, M. (1983). Crosslinguistic aspects of second language acquisition. *Applied Linguistics*, *4,* 192-199.

Thomas, E. R. (2011). *Sociophonetics: An introduction*. Basingstoke, Hampshire, New York, NY: Palgrave Macmillan.

Wester, F., Gilbers, D., & Lowie, W. (2007). Substitution of dental fricatives in English by Dutch L2 speakers. *Language Sciences*, *29*(2-3), 477-491.

To cite this article: Schmid, C., & Moosmüller, S. (2017). L2 acquisition and L1 attrition in Bosnian-German late bilinguals - Phonetic observations at the segmental level. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 251-257). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Vowel-to-vowel coarticulation in Greek normal-hearing and hearing-impaired speech

Anna Sfakianaki[1], Katerina Nicolaidis[2], Areti Okalidou[3]
anna.sfak@gmail.com, knicol@enl.auth.gr, okalidou@uom.edu.gr

[1, 2]Aristotle University of Thessaloniki, [3]University of Macedonia

Abstract. Speakers with prelingual profound hearing impairment (HI) develop speech without adequate auditory feedback. As a consequence, several aspects of their speech production are subject to delay or disorder. One particular aspect reported as deviant is coarticulation, the blending of articulatory gestures (Fowler, 1980) which ensures that speech is produced smoothly. Although acoustic research in other languages, mainly English, has focused on coarticulatory aspects in the productions of individuals with HI (i.e. Okalidou & Harris, 1999), for Greek this is an area that has only recently begun to be explored (Nicolaidis & Sfakianaki, 2016; Sfakianaki, 2012). Fourteen native speakers of Greek were recorded; five adults with normal hearing (NH) and nine adults with prelingual profound HI (>91 db HL) and with intelligible speech, using conventional hearing aids, uttered /pV_1CV_2/ disyllables with stress on the first or second syllable respectively, where V=/i, a, u/ and C=/p, t, s/. Magnitude and direction of V-to-V coarticulation was measured along the F1 and F2 at V1 offset (anticipatory) and V2 onset (carryover) across the three consonants. Findings showed that although both groups follow similar F1 and F2 coarticulation patterns, there are significant differences between the two groups in the degree of coarticulatory effects depending on consonantal context and coarticulatory direction. Overall more coarticulation was observed in alveolar contexts for speakers with HI. Variability in coarticulation is interpreted taking into account language-specific articulatory constraints. Results are discussed in light of speech production theories and the DAC model of articulatory constraints (Recasens & Espinosa, 2009), underscoring the importance of cross-linguistic research.

Keywords: hearing impairment, coarticulation, acoustic analysis, Greek

Introduction

Speech and hearing are interlinked. Loss of hearing limits the auditory input and feedback. If hearing loss occurs prelingually, that is before the speaker has managed to acquire speech and language, then speech production becomes challenging (Pratt, 2005). Although there is no ‘generic deaf speech pattern’ (McGarr & Harris, 1980:309), common problems observed in hearing impaired (henceforth HI) speech include vowel and consonant errors, as well as reduced phonological space due to lesser vowel differentiation arising from restricted tongue movement, primarily along the horizontal (F2) and secondarily along the vertical (F1) axis (e.g., Nicolaidis & Sfakianaki, 2016; Pratt & Tye-Murray, 2009; Ryalls, Larouche, & Giroux, 2003).

An area presenting difficulties for speakers with HI is correct timing in articulation. This has an impact on their gestural coordination in a variety of ways. In an attempt to explain gestural coordination and coarticulation in normal speech production various theories have been developed. According to the *Action Theory* or *Theory of Coproduction* (Fowler, 1980), phonemes are four-dimensional, articulatory entities that overlap temporally. A model developed within this framework is the *Degree of Articulatory Constraint (DAC) Model of Coarticulation* (Recasens, Pallarès, & Fontdevila, 1997). As maintained by the DAC model, the degree of tongue dorsum activation during the production of a phoneme is correlated with its coarticulatory resistance and aggression. As a consequence, in vowel-consonant-vowel (VCV) sequences, highly constrained segments in terms of tongue dorsum involvement induce large consonant-to-vowel coarticulatory effects and block vowel-to-consonant and vowel-to-vowel coarticulation. Previous research on coarticulation in HI speech has yielded contradictory findings. Early studies report less coarticulation and inconsistent coarticulatory patterns in HI speech (Monsen, 1976; Waldstein & Baum, 1991). However, more recent research revealed that speakers with HI may coarticulate less or more than speakers with normal hearing (henceforth NH) depending on the context (e.g., McCaffrey Morrison, 2008; Okalidou & Harris, 1999).

The present paper is a brief account of a larger investigation in the coarticulation of Greek NH and HI speech, and reports findings regarding vowel-to-vowel (V-to-V) F1 and F2 anticipatory and carryover coarticulatory effects in /pV_1CV_2/ disyllables comprising the three point vowels /i, a, u/ in the consonantal environment of the bilabial stop /p/, the alveolar stop /t/, and the alveolar fricative /s/. The choice of contexts was made on the basis of their difference in DAC value. This study tests the predictions of the DAC model with both NH and HI data from Greek. Coarticulatory patterns in NH vs. HI speech are investigated and discussed taking into account language-specific constraints, such as vowel inventory and phonemic vowel density (Manuel & Krakow, 1984).

Method

Fourteen native speakers of Greek participated in the experiment. The control (NH) group comprised five adults, two male and three female, aged 18-21. They were undergraduate students at the Aristotle University of Thessaloniki, with no reported hearing problems. The experimental (HI) group consisted of nine adults with prelingual profound hearing loss, five male and four female, aged 20-35. Most of them were university students. Certain criteria were set for the selection of speakers with HI: i) their hearing loss was prelinguistic, bilateral, stable and profound (91-105 dB Hz at 500, 1000 and 2000 Hz), ii) they had no cochlear implants, iii) they had been fitted with hearing aids before the age of 4 and wore them continuously, iv) they had received speech-language therapy from an early age and for over 7 years, v) they used primarily oral communication and had attended mainstream primary and secondary schools, vi) they all had intelligible speech.

The speakers uttered disyllables of the type /pV_1CV_2/, where V=/i, a, u/ and C=/p, t, s/ with stress on the first or second syllable respectively, embedded in the phrase: leje _____ pali 'say ____ again'. A list of 540 randomized phrases in total was read at a comfortable speaking rate and recorded at a sampling frequency of 22050 Hz. The data was edited and analysed with Praat (Boersma & Weenink, 2001). Any non target productions were excluded from the analysis. An LPC analysis measuring F1 and F2 formant frequencies (at V_1 offset for anticipatory effects and V_2 onset for carryover effects) was conducted using a 15ms Gaussian window. A total of 30,240 measurements were performed. Formant measurement is quite challenging especially in HI speech (see Sfakianaki & Nicolaidis, 2016). In order to examine V-to-V coarticulatory effects, disyllables were paired so that one member of the pair was always symmetrical and provided a 'fixed' context for the comparison (Recasens & Pallarès, 2000). In order to compare coarticulation effects between groups, two coarticulation variables, ΔF1 and ΔF2, were computed (e.g., in order to investigate /a/-to-/i/ carryover coarticulation over /s/: $\Delta F1 = F1_{\text{of /i/ in /pasi/}} - F1_{\text{of 2nd /i/ in /pisi/}}$ and $\Delta F2 = F2_{\text{of 2nd /i/ in /pisi/}} - F2_{\text{of /i/ in /pasi/}}$). Hence coarticulation (ΔF) was calculated based on within-subject formant differences. Univariate ANOVAs were performed for variables F1 and F2 (factors: hearing, measured vowel, transconsonantal vowel, consonant, stress, vowel position) and coarticulation variables ΔF1 and ΔF2 (factors: hearing, vowel pair, consonant, stress, direction) in SPSS (v. 17). Tukey pairwise post-hoc tests were carried out in Minitab (v. 15) in order to determine statistically significant within- and between-group differences.

Results

Regarding variable F1, statistical analyses showed that all factors were significant at both vowel offset and onset, while their interaction was found significant at vowel offset. ANOVAs for variable ΔF1 also revealed that all factors were statistically significant at vowel offset and onset, as well as their interactions at both vowel points. Regarding variable F2, all factors were also found significant at vowel offset. Their interaction was found significant at vowel onset. As far as variable ΔF2 is concerned, hearing was found significant at vowel onset and all remaining factors were found significant at both vowel offset and onset. Factor interaction at vowel offset was found statistically significant.

Table 1 presents the mean F1 and F2 difference (ΔF1 and ΔF2) and Standard Deviation (StDev) in Hz of the three fixed vowels, i.e., /i/, /a/ and /u/, according to vocalic and consonantal context for the NH and HI group. For this paper, data has been averaged across stress conditions due to space limitations. Tukey post-hoc tests were carried out in order to find out statistically significant differences among contexts within each group (asterisks in table 1, [*]: $p < .05$ or [**]: $p < .0001$), as well as between groups (crosses in table 1, [$^{+}$]: $p<.05$, [$^{++}$]: $p<.01$ or [$^{+++}$]: $p<.0001$).

Table 1. V-to-V anticipatory (at vowel offset) and carryover (at vowel onset) coarticulatory effects in the fixed vowel context of /i, a, u/ over the consonants /p, t, s/

	ΔF1 (StDev)				ΔF2 (StDev)			
Direction	Anticipatory		Carryover		Anticipatory		Carryover	
Group	NH	HI	NH	HI	NH	HI	NH	HI
Fixed /i/: /a/-to-/i/								
/p/	13 (41)	10 (51)	**56 (43)	7 (50)	91 (164)	**151 (230)	*102 (141)	68 (244)
/t/	17 (48)	-5 (49)	25 (43)	6 (48)	4 (120)	23 (150)	89 (89)	39 (113)
/s/	-3 (37)	-1 (38)	36 (94)	8 (39)	2 (124)	57 (143)	69 (114)	25 (113)
Fixed /i/: /u/-to-/i/								
/p/	8 (40)	3 (48)	25 (35)	6 (41)	**199 (157)	$^{+*}$98 (240)	**292 (161)	**109 (182)
/t/	5 (44)	-6 (45)	0 (42)	3 (42)	33 (113)	*93 (193)	*89 (98)	25 (97)
/s/	-3 (45)	-7 (41)	-5 (36)	0 (37)	60 (112)	*95 (179)	82 (92)	**91 (155)
Fixed /a/: /i/-to-/a/								
/p/	4 (58)	16 (85)	**52 (57)	*33 (66)	64 (94)	83 (126)	*87 (105)	**103 (146)
/t/	34 (58)	**46 (80)	**48 (55)	22 (60)	91 (79)	69 (107)	**171 (81)	**83 (104)
/s/	42 (117)	23 (83)	**50 (64)	**39 (69)	81 (103)	76 (132)	**208 (101)	61 (108)
Fixed /a/: /u/-to-/a/								
/p/	49 (73)	**55 (92)	**92 (61)	$^{+++**}$49 (72)	28 (94)	14 (95)	96 (102)	-2 (111)
/t/	*54 (64)	**71 (93)	**62 (58)	$^{+**}$32 (59)	16 (71)	90 (149)	-8 (88)	1 (105)
/s/	**78 (113)	**18 (168)	**70 (69)	**53 (68)	61 (102)	83 (175)	-61 (117)	19 (127)
Fixed /u/: /a/-to-/u/								
/p/	**60 (88)	21 (56)	**65 (71)	22 (64)	23 (138)	39 (134)	69 (101)	-27 (136)
/t/	0 (60)	6 (60)	33 (50)	-1 (41)	6 (135)	*104 (169)	49 (115)	-16 (156)
/s/	9 (50)	1 (44)	19 (49)	7 (43)	91 (259)	**94 (239)	-76 (164)	14 (205)
Fixed /u/: /i/-to-/u/								
/p/	**73 (98)	8 (54)	**50 (65)	14 (58)	69 (128)	34 (138)	**127 (138)	31 (145)
/t/	-13 (48)	-3 (62)	-2 (51)	-12 (39)	52 (138)	**137 (232)	**181 (143)	53 (148)
/s/	-10 (46)	-3 (41)	-3 (49)	-4 (41)	112 (238)	**93 (249)	*99 (184)	63 (206)

Concerning coarticulation on the F1 axis, we observe that both groups follow similar coarticulation patterns overall. The fixed /a/ context shows more effects for both groups, while the high vowel contexts /i/ and /u/ appear more constrained, especially for the HI group, and allow for effects only across the bilabial for the NH group. Although both groups demonstrate effects in the fixed /a/ context, the NH group displays significantly greater coarticulatory magnitude especially in the carryover direction. On the F2 axis, both groups show significant coarticulation in a number of contexts, although effects for the speakers with NH are greater in magnitude. For both groups, coarticulation on /a/ and /u/ seems to be greater across the alveolars, while for the NH group more effects are allowed across the bilabial in the fixed /i/ context. The two groups seem to favour different coarticulatory directions, with the NH group displaying more carryover and the HI group more anticipatory effects.

Discussion

The DAC model makes predictions about V-to-V effects, but *both* the constraints of the intervocalic consonant *and* that of the fixed vowel need to be taken into account (Recasens *et al*., 1997). Hence, less V-to-V effects are expected in the fixed front /i/ context than in the context of the central /a/ and possibly the back /u/, while more transconsonantal effects are expected to be allowed over the bilabial

/p/ than the alveolars /t/ and /s/. Between the two alveolars, smaller effects are expected across /s/ than /t/ due to the high articulatory and aerodynamic constraints involved in its production.

Regarding F1 V-to-V coarticulation for both groups, more coarticulation appears on the open vowel /a/ than the high vowels /i/ and /u/ in accordance with the predictions of the DAC model and research in other languages (e.g., Recasens & Pallarès, 2000, for Catalan; Mok, 2011, for Thai). In agreement with the DAC model, the NH high vowels /i/ and /u/ receive significant lowering V-to-V effects across the bilabial, while effects over the two alveolars are essentially blocked. In HI speech, no effects are located on the high vowels regardless of consonantal context. Hence, the high vowels of speakers with HI seem to be more constrained in height than those of speakers with NH. It is noteworthy that, regarding the fixed /a/, the HI group displays significant V-to-V effects in as many or even more contexts than the NH group, and post-hoc tests reveal no statistical difference in coarticulatory magnitude in the majority of the contexts between the two groups. Another interesting observation concerns the appearance of significant coarticulation across the alveolars for both groups, and especially across /s/ for the NH group, which, in some fixed-vowel contexts, surpasses the effects across the bilabial /p/. This finding is contrary to the predictions of the DAC model, but it could be associated with language-specific constraints. For example, the NH Greek /s/ has been found less constrained than English /s/, as tongue positioning and groove width are less critical for Greek /s/ due to lack of other contrastive fricatives in the alveolar region (Nicolaidis, 1997). As far as directionality is concerned, NH coarticulation appears more prominent in the carryover direction, a finding associated with the slow articulation of the massive jaw (Recasens & Pallarès, 2000; Sussman, MacNeilage, & Hanson, 1973), while this preference is not as clear in HI coarticulation.

Regarding V-to-V F2 coarticulation for both groups, effects between the two high vowels, i.e., /i/-to-/u/ and /u/-to-/i/, are more pronounced than between /a/ and the high vowels, as expected due to F2 distance. In accordance with the DAC model and evidence on other languages (e.g., Gay, 1977, for American English; Recasens & Espinosa, 2009 for Catalan; Mok, 2011 for Thai), the NH front vowel /i/ exhibits more coarticulatory aggression than the vowels /a/ and /u/. However, /i/ does not exhibit the expected coarticulatory resistance, as it undergoes substantial backing especially from /u/ in the bilabial context. Hence, in agreement with previous studies on Greek NH speech (Nicolaidis, 1997), Greek /i/ displays more variability than that predicted by the DAC model for a front vowel. This result is in line with Manuel and Krakow's (1994) output constraint hypothesis, according to which, vowels of languages with a smaller inventory size may be more coarticulated than vowels of languages with denser vocalic systems. In HI speech, the two alveolars allow for increased anticipatory V-to-V effects in the fixed high vowel context compared with NH speech. Similar results for anticipatory V-to-V coarticulation effects in the alveolar /d/ context have been reported for American deaf speakers (Okalidou & Harris, 1999). Previous research has shown that the fricative /s/ is more constrained overall in HI speech, as speakers with HI may produce a simultaneous postalveolar and medio/postpalatal constriction for /s/ (Nicolaidis, 2004). When a highly constrained alveolar consonant is followed by a high vowel, the whole CV_2 syllable tends to become palatalized. Hence, such CV syllables are more overlapped for speakers with HI and may, therefore, induce stronger V-to-V anticipatory effects whilst blocking V-to-V carryover effects. CV syllables tend to be largely overlapped in early child speech, especially when both the consonant and the vowel involve the same articulatory subsystem (Goodell & Studdert-Kennedy, 1993), and thus HI speech can be viewed as similar to developing speech in this respect. Finally, the two groups present a more similar F2 V-to-V coarticulatory pattern in the fixed low mid /a/ context, showing significant effects only in the carryover direction. In accordance with the DAC model, significant effects are induced only from /i/ due to the longer articulatory distance involved.

Conclusion

In sum, comparing HI high vowels /i/ and /u/ with the corresponding NH vowels, it is seen that they are more susceptible to influences from the transconsonantal vowel in tongue front/back displacement than from jaw/tongue height movement. The open vowel /a/ seems almost equally susceptible in jaw/tongue height displacement for both groups but relatively less influenced in tongue front/back movement in the carryover direction for the HI than the NH group. Hence, our results show that, depending on fixed vowel context, HI speech shows more or less transconsonantal influence than that found in NH speech. The majority of the findings are well accounted for by the DAC model of coarticulation (Recasens *et al*., 1997) based on the coproduction framework. Our data suggest that HI V-to-V coarticulation patterns share more similarities to the NH patterns on the F1 than the F2 axis. An interesting finding of the present study concerns coarticulatory direction. V-to-V coarticulation across alveolar consonants was found more prominent in HI than NH speech in the anticipatory direction, especially in the F2 formant frequency. This result may be related to increased palatalization during alveolar production in HI speech requiring stronger anticipation of the whole CV_2 syllable and is in line with Okalidou & Harris (1999), who also found more intervocalic anticipatory coarticulation in the alveolar context in HI than NH American English speech.

Overall, our findings suggest a more constrained alveolar production and a more overlapped CV_2 syllable when both the consonantal and the vocalic gestures involve tongue-dorsum activation, denoting less control of the tongue articulators by speakers with HI. A concurrent investigation of consonant-to-vowel coarticulatory effects on both F1 and F2 axes might illuminate aspects of interarticulatory organization in HI speech and further explicate similarities and differences between NH and HI coarticulation patterns (Sfakianaki, Nicolaidis, & Okalidou, in preparation). Finally, such findings can bear important implications for clinical practice. The differences noted between the two groups suggest that speakers with HI may benefit from speech therapy protocols that focus on the coproduction of gestures in sequences containing different lingual consonants and vowels, so as to enhance spatio-temporal coordination in a variety of VCV sequences.

References

Boersma, P., & Weenink, D. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341-345.

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing. *Journal of Phonetics, 8,* 113-133.

Gay, T. (1977). Cinefluorographic and electromyographic studies of articulatory organization. In M. Sawashima & F. S. Cooper (eds.), *Dynamic aspects of speech production* (pp. 85-102). Tokyo: University of Tokyo Press. February 10, 2012 Online: http://www.haskins.yale.edu/Reprints/HL0227.pdf

Goodell, E. W., & Studdert-Kennedy, M. (1993). Acoustic evidence for the development of gestural coordination in the speech of 2-year-olds: A longitudinal study. *Journal of Speech and Hearing Research, 36,* 707-727.

Manuel, S., & Krakow, R. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Status Report on Speech Research*, *SR-77/78*, 69-78. Haskins Laboratories.

McCaffrey Morrison, H. (2008). The locus equation as an index of coarticulation in syllables produced by speakers with profound hearing loss. *Clinical Linguistics and Phonetics, 22*(9), 726-740.

McGarr, N. S., & Harris, K. (1980). Articulatory control in a deaf speaker. *Status Report on Speech Research, SR-63/64* (pp. 309-332). Haskins Laboratories.

Mok, P. K. P. (2011). Effects of vowel duration and vowel quality on vowel-to-vowel coarticulation. *Language and Speech*, 0023830911404961, first published on June 24, 2011 doi:10.1177/0023830911404961

Monsen, R. B. (1976). Second formant transitions of selected consonant-vowel combinations in the speech of deaf and normally-hearing children. *Journal of Speech and Hearing Research, 19,* 279-289.

Nicolaidis, K. (1997). *Sources of articulatory variability in Greek: an electropalatographic study*. Unpublished Doctoral Dissertation. University of Reading.

Nicolaidis, K. (2004). Articulatory variability during consonant production by Greek speakers with hearing impairment: an electropalatographic study. *Clinical Linguistics & Phonetics*, *18*(6-8), 419-432.

Nicolaidis, K., & Sfakianaki, A. (2016). Acoustic characteristics of vowels produced by Greek intelligible speakers with profound hearing impairment II: The influence of stress and context. *International Journal of Speech-Language Pathology*, *18*(4), 378-387.

Okalidou, A., & Harris, K. S. (1999). A comparison of intergestural patterns in deaf and hearing adult speakers: Implications from an acoustic analysis of disyllables. *Journal of the Acoustical Society of America, 106(1),* 394-410.

Pratt, S. R. (2005). Aural habilitation update: The role of auditory feedback on speech production skills of infants and children with hearing loss. *ASHA Leader, 10*(4), 8-9, 32-33.

Pratt, S. R., & Tye-Murray, N. (2009). Speech impairment secondary to hearing loss. In M. R. McNeil (ed.), *Clinical management of sensorimotor speech disorders* (2nd edn.) (pp. 204-234). New York, NY: Thieme Medical Publishers.

Recasens, D., & Espinosa, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *Journal of the Acoustical Society of America*, *125(4),* 2288-2298.

Recasens, D., & Pallarès, M. D. (2000). A study of F1 coarticulation in VCV sequences. *Journal of Speech, Language, and Hearing Research*, *43,* 501-512.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *Journal of the Acoustical Society of America*, *102*(1), 544-561.

Ryalls, J., Larouche, A., & Giroux, F. (2003). Acoustic comparison of CV syllables in French-speaking children with normal hearing, moderate-to-severe and profound hearing impairment. *Journal* of *Multilingual Communication Disorders*, *1*(2), 99-114.

Sfakianaki, A. (2012). F1 coarticulation phenomena in the productions of Greek speakers with hearing impairment. In Z. Gavriilidou, A. Efthymiou, E. Thomadaki, & P. Kambakis-Vougiouklis (eds), *Selected papers of the 10th International Conference of Greek Linguistics* (pp. 520-527). Komotini/Greece: Democritus University of Thrace. September 15, 2017. Online: http://www.icgl.gr/files/English/45.Sfakianaki_10ICGL_pp.520-527.pdf

Sfakianaki, A., & Nicolaidis, K. (2016). Acoustic aspects of segmental and suprasegmental productions of Greek hearing-impaired speech: A qualitative analysis. In M. Mattheoudakis & K. Nicolaidis (eds.), *Selected Papers of the 21st International Symposium on Theoretical and Applied Linguistics (ISTAL 21)* (pp. 401-425), ISSN 2529-1114. Aristotle University of Thessaloniki. September 15, 2017. Online: https://ejournals.lib.auth.gr/thal/article/view/5239/5125

Sussman, H. M., MacNeilage, P. F., & Hanson, R. J. (1973). Labial and mandibular dynamics during the production of bilabial consonants: Preliminary observations. *Journal of Speech and Hearing Research*, *16,* 397-420.

Waldstein, R. S., & Baum, S. R. (1991). Anticipatory coarticulation in the speech of profoundly hearing-impaired and normally hearing children. *Journal of Speech & Hearing Research*, *34,* 1276-1285.

To cite this article: Sfakianaki, A., Nicolaidis, K., & Okalidou, A. (2017). Vowel-to-vowel coarticulation in Greek normal-hearing and hearing-impaired speech. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 258-263). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Verbal short term memory and receptive language abilities in bilingual children with specific language impairment

Ioanna Talli, Stavroula Stavrakaki
tallijoan@yahoo.com, svoula@itl.auth.gr

Aristotle University of Thessaloniki

Abstract. The present study investigates whether bilingual children with specific language impairment (SLI) show advantages in verbal short term memory (STM) and, if so, whether it affects lexical and syntactic abilities. Sixteen monolingual and sixteen bilingual children with SLI with Greek as L1 and Albanian or Russian as L2 were compared with two control groups of 20 monolingual and 18 bilingual (L1: Albanian or Russian, L2: Greek) typically developing (TD) children in terms of vocabulary, verbal STM and syntactic comprehension. The results indicate parallel patterns of deficits in bilingual and monolingual children with SLI in the domains of STM and syntax.

Keywords: specific language impairment, verbal STM, syntax, bilingualism

Introduction

Children with specific language impairment (SLI) show poor oral language development that cannot be attributed to neurological, sensorimotor, mental or emotional/psychological deficits (Bishop, 1997; Leonard 1998/2014). Cross-linguistic findings have revealed that difficulties in complex syntax and phonological short term memory (STM) are particularly evident in children with SLI (Adani, Forgiarini, Guasti, & van der Lely 2014; Archibald & Gathercole, 2006; Conti-Ramsden, 2003; Friedmann, Belletti, & Rizzi 2009; Lalioti, Stavrakaki, Manouilidou, & Talli, 2016; Newbury, Bishop, & Monaco, 2005; Stavrakaki & van der Lely, 2010).

Recent studies examined whether exposure to two languages deteriorates symptoms of SLI (for a review, see Paradis, 2010). Paradis and colleagues conducted many studies on French-English speaking bilingual children with SLI (BI-SLI) simultaneously exposed to two languages and found no negative effect of this exposure on their linguistic development (e.g., Paradis, Crago, & Genesee, 2005/2006; Paradis, Crago, Genesee, & Rice, 2003). Other studies tried to identify whether typically developing children (TD) successively exposed to a second language (L2) showed similar traits in their linguistic development, to those of monolingual children with SLI. Some of these studies concluded that monolingual children with SLI show a distinct linguistic profile and qualitative differences compared to TD children exposed to an L2 during language development (Clahsen, Rothweiler, Sterner, & Chila, 2014; Rothweiler, Chila, & Clahsen, 2012, among others) while other studies identified similar traits in monolingual SLI and typical bilingual acquisition (e.g., Haiden, Ferré, Prévost, Scheidnes, & Tuller, 2009).

More recently, studies of BI-SLI have explored whether verbal memory is impaired in these children or, by contrast, whether it is beneficially affected in bilingual language impairment. Previous studies of typically developing bilingual children (BI-TD) revealed advantages in verbal STM (Morales Calvo, & Bialystok, 2013; Soliman, 2014), and thus, it is interesting to examine whether this advantage holds for children with BI-SLI. Blom and Boerma (2017) have examined the effects of bilingualism on language impairment (LI) by comparing monolingual and bilingual TD and LI children in verbal memory tasks, as well as in vocabulary and morphology. They suggested that while vocabulary and morphology were negatively affected by bilingualism, verbal memory was beneficially affected. Furthermore, Meir (2017) extensively tested Hebrew-Russian bilingual children with SLI on a range of verbal STM tasks. She found no beneficial effects of bilingualism *per se* on SLI performance as participants with SLI showed deficient performance on all experimental tasks, tapping into verbal STM. By contrast, in BI-TD negative effects of bilingualism were only observed when the linguistic load of the verbal STM task was increased (sentence repetition). She attributes this performance to the small vocabulary sizes in BI-TD.

The present study aims at contributing to the current discussion on the effect that bilingualism has on the linguistic and cognitive abilities of children with SLI by testing Russian-Greek and Albanian-Greek children with SLI on syntactic and verbal short-term memory tasks and comparing their performance to that of TD monolingual and Russian-Greek, Albanian-Greek bilingual children, as well as to that of monolingual Greek children with SLI. In addition to testing vocabulary and verbal STM abilities, we explore the receptive syntactic abilities of these children and investigate possible interrelation between these linguistic and cognitive domains.

Method

Participants

There were four groups of children in total, two clinical groups and two control groups:

i) The clinical groups included 16 monolingual and 16 bilingual children with SLI, with Albanian or Russian as a first language (L1) and Greek as a second language (L2). These children were diagnosed with SLI by experienced speech and language therapists working in public or private centers for speech and language therapy in Greece. All these primary school aged children were reported to show language delay or disorders at preschool age and continued to show persistent language deficits at the time of testing. As far as the BI-SLI children are concerned, all of them were born in Russian or Albanian speaking families, but most of them had been exposed to Greek since birth as their parents spoke Greek at home. Specifically, only two of these children were exposed to Greek at the age of 4 in kindergarten. Greek is the dominant language for all of them at the time of testing. All participants with SLI were pre-experimentally tested by means of the Greek version of RAVEN Standard Progressive Matrices (Raven, 1947) and showed normal IQ abilities. Furthermore, their vocabulary abilities were assessed using the Peabody Picture Vocabulary Test (PPVT) (Dunn & Dunn 1981; Greek adaptation: Simos, Sideridis, Protopapas, & Mouzaki, 2011).

ii) The two control groups included 20 monolingual Greek and 18 bilingual (Greek as L1) TD children, respectively. These children were matched to the two clinical groups in chronological age and non-verbal IQ (Raven's Coloured Progressive Matrices). The control groups did not differ significantly in chronological age and in Raven percentile from the two clinical groups. Information on participant groups is shown in Table 1 below.

Table 1. Group profiles

Groups	Number	Number of females	CA Mean (SD)	Non-verbal IQ percentile (measured by RAVEN) Mean (SD)	Vocabulary percentile (measured by PPVT) Mean (SD)
MONO-SLI	16	5	8;1 (1;5)	59.3 (20.14)	24.9 (19.2)
BI-SLI	16	4	8;4 (1;4)	51.2 (20.7)	9.9 (12.8)
MONO-CONTROLS	20	5	9 (1;4)	67.8 (20.9)	65.8 (19.8)
BI-CONTROLS	18	6	8;6 (1;8)	56.39 (20.3)	39.2 (23.3)

One-way ANOVA showed that groups did not significantly differ in non-verbal IQ as measured by the Raven test, despite the apparently higher performance of typically developing monolingual children, as shown in Table 1 (F (3, 66)= 2.282, p = .087). By contrast, as far as the vocabulary percentiles (PPVT) are concerned, significant between group differences were found (F (3, 69) = 27.5, p <.001) with BI-SLI showing the lowest performance, while the BI-controls and MONO-SLI children showed almost the same performance. Specifically, bilingual children with SLI had significantly lower performance than all groups: MONO-SLI (t (30) = 2.60, p<.05), MONO-TD (t

(34) = -9.78, *p* <.001) and BI-TD (*t* (32) = -4.26, *p* < .001). However, the difference between MONO-SLI and BI-TD was not significant (*t* (32) = -1.94, *p* = .61), while MONO-TD children had significantly better performance than BI-TD children (*t* (36) = 3.81, *p* = .001).

Tasks and procedure

In this study, we employed tasks to assess verbal STM as well tasks to assess syntactic comprehension. The experimental tasks were:

Verbal STM

Non-word repetition. This test adapted in Greek from the French battery EVALEC (Sprenger-Charolles, Colé, Béchennec, & Kipffer Piquard, 2005; Greek adaption: Talli, 2010) included the repetition of 24 three- to six-syllable non-words (6 non-words for each length) presented in order of increasing length. The children had to repeat each item the examiner presented orally. The total number of syllables correctly repeated was the accuracy score, calculated in percentage.

Digit span. The Digit Span subtest of Greek WISC-III (Georgas, Paraskevopoulos, Besevegis, & Giannitsas, 1997) consists of two tasks, a forward and a backward digit span task. It requires that the child repeats first forward and then backwards a list of digits of increasing length, orally presented by the examiner. We calculated the total raw scores of both subtests.

Sentence repetition. This test is a subtest of the *Diagnostic Test of Verbal Intelligence* (DVIQ, Stavrakaki & Tsimpli, 2000). It consists of 15 sentences with a maximum score of 45. The children had to repeat each sentence the examiner presented orally. We calculated the total correct score in percentages.

Syntactic comprehension

This task consisted of 36 sentences in total (24 relative clauses including subject and object gap relatives, 4 with reflexive verbs, and 8 with passive voice). Percentages of accuracy scores were calculated for all types of sentences. Each child was tested individually. The examiner presented each sentence orally, while the child was shown a page with four choices of pictures, and had to select the picture that matched the spoken sentence.

In Table 3, examples of all experimental sentences are presented.

Table 3. Examples of experimental sentences

subject relatives	e.g., *η κυρία που φιλάει τον κύριο σπρώχνει την κοπέλα* 'the lady that is kissing the man is pushing the girl'
object relatives	e.g., *η κοπέλα σπρώχνει τον κύριο που κρατά η νοσοκόμα* 'the young lady is pushing the man that the nurse is holding'
passive sentences	e.g., *το αγόρι κλωτσιέται από το κορίτσι* 'the boy is kicked by the girl'
reflexive verbs	e.g., *το κορίτσι λούζεται στο μπάνιο* 'the girl is washing herself in the bathroom'

Results

Results are presented in Table 4 below. Both clinical groups performed below TDC in all tasks. In addition, BI-SLI showed comparable performance with MONO-SLI in most of the STM tasks while they performed below MONO-SLI in syntactic tasks. With respect to TD, BI-TD performed below MONO-TD in STM tasks and above MONO-TD in syntactic tasks, but these performances were comparable. To investigate whether these observations were valid, we performed statistical analysis whose results are summarized in Table 5. Specifically, for the different tasks, ANOVAs were conducted with the 4 groups as the between subject factor. When the effect of group was significant, T-tests were conducted to compare performances of MONO-SLI and BI-SLI with MONO-TD, between MONO-SLI and BI-SLI with BI-TD and between the two clinical groups.

For sentence repetition, there was a main effect of group (*F* (3, 69) = 7.53, *p*<.001). The difference between the two clinical groups just failed to be significant (*t* (30) = 1.98, *p* = .06). MONO-SLI were worse than MONO-TD children (*t* (34) = -3.37, *p* <.05), but not worse than BI-TD children (*t* (32) = 1.22, *p* = .23). Therefore, the same performance level was shown by monolingual children with SLI and BI-TD children on sentence repetition. BI-SLI children were worse than MONO-TD (t(34) = -3.54, p=.001) and BI-TD children (*t* (32) = -2.67, *p*<.05). Finally, MONO-TD children had significantly better performance than BI-TD children (t (36) = 2.78, *p* < .05).

Table 4. Mean performance (and SDs) of groups on experimental tasks (all in percentages)

	Mono-SLI	Bi-SLI	Mono-TD	Bi-TD
STM tasks				
Digit span	32.08 (9.3)	28.13 (6.7)	41.5 (8.2)	36.1 (6.8)
Non-word repetition	77.6 (12.9)	70.6 (9.7)	91.6 (5.8)	84.2 (8.3)
Sentence repetition	95.4 (5.4)	87.4 (15.4)	99.6 (1.2)	97.3 (3.5)
Syntactic tasks				
Relative clauses	67.5 (16,4)	57.5 (143)	81.5 (10.4)	84.4 (5.7)
Passives	66.4 (26,5)	45.3 (254)	79.4 (17.8)	89.6 (18.3)
Reflexive verbs	81.3 (26,6)	68.8 (19.4)	93.7 (13.8)	98.6 (5.9)

With regard to non-word repetition, the group effect was significant (*F* (3, 69) = 16.62, *p*<.001). The difference between the two clinical groups was not significant (*t* (30) = 1.74, *p* = .09). MONO-SLI were worse than MONO-TD children (*t* (34) = -4.36, *p* < .001), but they just failed to perform significantly worse than BI-TD children (*t* (32) = -1.79, *p*=.08). BI-SLI children were worse than MONO-TD (*t* (34) = -8.06, *p* < .001) and BI-TD children (t (32) = -4.40, *p* < .001). Finally, MONO-TD children had significantly better performance than BI-TD children (*t* (36) = 3.21, *p* < .05).

With regard to the digit span, there was a main effect of group (*F* (3, 69) = 9.64, *p*<.001). The difference between the two clinical groups was not significant (t (30) = 1.39, p = .18). MONO-SLI were worse than MONO-TD children (*t* (34) = -3.23, *p*<.05), but not worse than BI-TD children (*t* (32) = -1.46, *p*=.16). Therefore, the same performance level was shown in monolingual children with SLI and BI-TD children on digit span. BI-SLI children were worse than MONO-TD (*t* (34) = -5.28, *p* <.001) and BI-TD children (*t* (32) = -3.45, *p* <.05). Finally, MONO-TD children had significantly better performance than BI-TD children (*t* (36) = 2.19, *p* < .05).

With regard to the relative clauses, one-way ANOVA indicated a main effect of group (F(3, 69) = 18.40, p<.001). The difference between the two clinical groups was not significant (t (30) = 1.85, p = .07). MONO-SLI were worse than MONO-TD children (t (34) = -3.11, p<.05), and worse than BI-TD children (*t* (32) = -4.09, *p* <.001). Interestingly, in this task MONO-SLI performed lower than BI-TD. BI-SLI children were worse than MONO-TD (*t* (34) = -5. 86, *p* < .001) and BI-TD children (*t* (32) = -7.39, *p* < .001). Finally, MONO-TD children had comparable performance with BI-TD children (*t* (36) = -.05, *p* = .30).

With regard to passives, the group effect was significant (*F* (3, 69) = 12.78, *p* < .001). The difference between the two clinical groups was significant and BI-SLI had worse performance than MONO-SLI (*t* (30) = 2.30, *p* = .03). Notably, in this syntactic task, BI-SLI had worse performance than MONO-SLI. MONO-SLI had comparable performance with MONO-TD children (*t* (34) = -1.75, p=.09), but were significantly worse than BI-TD children (*t* (32) = -2.99, p<.05). BI-SLI children were worse than MONO-TD (*t* (34) = -4.73, *p* < .001) and BI-TD children (*t* (32) = -5.89, *p* < .001). Finally, MONO-TD children had comparable performance with BI-TD children (*t* (36) = -1.74, *p* = .09).

With regard to reflexive verbs, there was a main effect of group (*F* (3, 69) = 9.88, *p* <.001). The difference between the two clinical groups was not significant (*t* (30) = 1.52, *p* = .14). MONO-SLI had comparable performance with MONO-TD children (*t* (34) = -1.82, *p*=.08), but were significantly worse than BI-TD children (*t* (32) = -2.70, p < .05). BI-SLI children were worse than MONO-TD (*t* (34) = -4.53, *p* < .001) and BI-TD children (*t* (32) = -6.24, *p* < .001). Finally, MONO-TD children had significantly better performance than BI-TD children (*t* (36) = -1.39, *p* = .17).

In sum, both clinical groups had significantly poorer performance compared to the monolingual TD group in all tasks (except for passives and reflexives for mono-SLI). BI-SLI children had significantly worse performance than mono-SLI children in one syntactic comprehension task (passives). Comparisons between BI-SLI and BI-TD showed that BI-SLI children were significantly worse than BI-TD children in all tasks. Furthermore, while no significant differences were attested between TD-MONO and TD-BI, the performance of TD-BI children was higher in syntactic tasks.

Table 5. Between group comparisons and statistically significant (or no statistically significant) group differences

	Digit span	Non-word repetition	Sentence repetition	Relative clauses	Passive voice	Reflexive verbs
MONO-SLI	=BI-SLI	=BI-SLI	=BI-SLI	=BI-SLI	>BI-SLI	=BI-SLI
	<MONO-TD	<MONO-TD	<MONO-TD	<MONO-TD	=MONO-TD	=MONO-TD
	=BI-TD	=BI-TD	=BI-TD	<BI-TD	<BI-TD	<BI-TD
BI-SLI	<MONO-TD	MONO-TD	<MONO-TD	<MONO-TD	<MONO-TD	<MONO-TD
	<BI-TD	<BI-TD	<BI-TD	<BI-TD	<BI-TD	<BI-TD
MONO-TD	>BI-TD	>BI-TD	>BI-TD	=BI-TD	=BI-TD	=BI-TD

< indicates statistically significantly below performance ($p < .05$)
= indicates no statistically significant differences ($p > .05$)

Discussion

The present study was strongly motivated by results indicating linguistic advantages for TD bilingual children (e.g., Morales *et al*., 2013; Soliman, 2014) and significant linguistic deficits in children with SLI in the domains of syntax and verbal short-term memory (Leonard, 1998/2014 for a review). The main question of this study was whether exposure to more than one languages could benefit bilingual children with SLI. For this purpose, we tested four groups of children on a range of verbal STM and syntactic tasks in addition to a vocabulary (clinical) task. The results indicated that deficits in receptive vocabulary were more striking in bilingual children with SLI than in monolingual ones. This was found to hold in typical development, as well: BI-TD participants performed below MONO-TD participants. Notably, in this task, the BI-TD participants performed almost like MONO-SLI. These results are in agreement with those reported by Blom and Boerma (2017) and Meir (2017) which show low vocabulary performance for TD bilingual children with SLI and thus confirm the difficulties in the vocabulary acquisition by bilingual children.

With respect to the verbal STM tasks, the BI-TD group performed below MONO-TD group in these tasks mediated by language (non-word repetition and sentence repetition). Notably, the performance of monolingual children with SLI was at the same level with that shown by typically developing bilingual children on STM tasks but significantly below on the reception of syntax. This shows that the receptive syntactic abilities of BI-TD participants were enhanced compared to their verbal STM abilities which were nevertheless mediated by expressive language. Apparently, these results do not confirm previous findings showing advantages in verbal STM for BI-TD children (Soliman, 2014) Notably, our results showed advanced syntactic abilities by the BI-TD group despite their limitations in vocabulary. Impressively, they performed higher than MONO-TD participants.

In addition, relative clauses appear to be a robust marker for MONO- and BI-SLI as they performed significantly below both TD groups. However, the other syntactic tasks did not appear to provide a clear-cut distinction between monolingual TD and SLI, contrary to previous findings reported for Greek (Stavrakaki, 2005, for a review). It may be the case that passive voice, due to its highly infrequent use in Greek, does not constitute a marker for the acquisition of syntax per se. In addition, as reflexive verbs are formed in the lexicon (Holton, Mackridge, & Philippaki-Warburton, 1997; Tsimpli, 1989), they may not also provide a crucial index for syntactic abilities.

In conclusion, the abovementioned results indicate that deficits in receptive vocabulary are more evident in BI-TD and BI-SLI participants than in their monolingual peers. Moreover, the verbal STM does not appear to be superior in children with BI-SLI or to affect their performance on syntax or on vocabulary. Our results showed that exposure to a second language did not seem to be a factor that facilitated BI-SLI children's performance on the reception of syntax and verbal STM. However, it was not shown to be burdensome as indicated by all, except one, comparisons with monolingual children with SLI. These findings highlight the severely impaired status of verbal STM and syntactic abilities in children with SLI independently of the number of languages these children are exposed to.

References

Adani, F., Forgiarini, M., Guasti, M. T., & van der Lely, H. K. (2014). Number dissimilarities facilitate the comprehension of relative clauses in children with (grammatical) specific language impairment. *Journal of Child Language*, *41*(4), 811-841.

Archibald, L., & Gathercole, S. (2006). Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research*, *49*, 970-983.

Blom, E., & Boerma, T. (2017). Effects of language impairment and bilingualism across domains. *Linguistic Approaches to Bilingualism*, *7*(3-4), 277-300.

Clahsen, H., Rothweiler, M., Sterner, F., & Chilla, S. (2014). Linguistic markers of specific language impairment in bilingual children: The case of verb morphology. *Clinical Linguistics and Phonetics*, *28*(9), 709-721.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment. *Journal of Speech, Language and Hearing Research*, *46*, 1029-1037.

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized Relatives. Types of intervention in the acquisition of A-bar dependencies. *Lingua*, *119*, 67-88.

Georgas, J., Paraskevopoulos, I. N., Besevegis, E., & Giannitsas, N. D. (1997). *The Hellenic WISC-III.* Athens: Psychometric Laboratory, University of Athens.

Haiden, M., Ferré, S., Prévost, P., Scheidnes, M., & Tuller, L. (2009). Production and Comprehension of Wh-questions in the acquisition of questions in the acquisition of French: Comparing L2 children and children with SLI. Presented at DGfS 2009, Universität Osnabrück, March 4th, 2009. Available at https://www.uni-frankfurt.de/49015381/Haiden et al1.pdf, accessed on the 10th of November 2017.

Holton, D., Mackridge, P., & Philippaki-Warburton, I (1997*). Greek: A Comprehensive Grammar of the Modern Language.* London, UK: Routledge

Lalioti, M., Stavrakaki, S., Manouilidou C., & Talli, I. (2016). Subject-verb agreement and verbal short term memory: A perspective from Greek children with specific language impairment. *First Language*, *36*, 279-294.

Leonard, B. L. (1998/2014). *Children with specific language impairment*. Cambridge, MA: MIT Press.

Meir, N. (2017). Effects of Specific Language Impairment (SLI) and bilingualism on verbal short-term memory. *Linguistic Approaches to Bilingualism*, 7(3-4), 301-330.

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*(2), 187-202.

Newbury, D. F., Bishop, D. V., & Monaco, A. P. (2005). Genetic influences on language impairment and phonological short-term memory. *Trends in Cognitive Sciences, 9*(11), 528-534.

Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics, 31*(2), 227-252.

Paradis, J., Crago, M., & Genesee, F. (2005/2006). Domain-specific versus domain-general theories of the deficit in SLI: Object pronoun acquisition by French–English bilingual children. *Language Acquisition*, *13*, 33-62.

Paradis, J., Crago, M., Genesee, F., & Rice, M. (2003). Bilingual children with specific language impairment: How do they compare with their monolingual peers? *Journal of Speech, Language, and Hearing Research, 46*, 1-15.

Rothweiler, M., Chilla, S., & Clahsen, H. (2012). Subject–verb agreement in specific language impairment: A study of monolingual and bilingual German-speaking children. *Bilingualism: Language and Cognition*, *15*(1), 39-57.

Simos, P., Sideridis, G.D., Protopapas, A., & Mouzaki, A. (2011). Psychometric characteristics of a receptive vocabulary test for Greek elementary students. *Assessment for Effective Intervention*, *37*, 34-50.

Soliman, A. M. (2014). Bilingual advantages of working memory revisited: A latent variable examination. *Learning and Individual Differences*, 32, 168-177.

Stavrakaki, S. (2005).). Greek neurolinguistics: The state-of-the-art. *Journal of Greek Linguistics*, *5*, 181-228.

Stavrakaki, S., & van der Lely, H. (2010). Production and comprehension of pronouns by Greek children with specific language impairment. *British Journal of Developmental Psychology*, *28*(1), 189-216.

Talli, I. (2010). *Linguistic abilities in developmental dyslexia and specific language impairment: A comparative and cross-linguistic approach*. Unpublished PhD thesis. Paris Descartes University and Aristotle University of Thessaloniki.

Tsimpli, I. M. (1989). On the properties of the passive affix in Modern Greek. *UCL Working Papers in Linguistics*, *1*, 236-261.

To cite this article: Talli, I., & Stavrakaki, S. (2017). Verbal short term memory and receptive language abilities in bilingual children with specific language impairment. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 264-270). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Translanguaging: A synthesis of recent literature and suggestions for future directions of research and action

Emmanouela Tisizi
emmanouela.tisizi@mail.mcgill.ca

McGill University

Abstract. In a world characterized by globalization and people movement, now more than ever, classes of students from diverse cultural, ethnic and linguistic backgrounds are formed, and teachers are required to change their methods so as to create safe and inclusive teaching environments. This cannot be achieved, however, if students' various linguistic repertoires continue to be treated as impediments to the learning practice. Indeed, for many years, bi/multilinguals have been viewed as the sum of two monolinguals coexisting in the same person and their language skills have often been evaluated in monolingual terms (Grosjean, 2010). However, bi/multilinguals actually have a very complex linguistic configuration and their languages interact constantly and create a unique language system (Grosjean & Li, 2013; Pavlenko, 2014). It is essential to recognize the value of bi/multilinguals' prior knowledge and build on it to create new knowledge. Teachers have the responsibility to 'teach the whole child' (Cummins, Bismilla, Chow, Cohen, Giampapa, Leoni, Sandhu, & Sastri, 2005). That is, students' existing knowledge (both cultural and linguistic) must be used as a *resource* to enable the development of new knowledge. This can facilitate the acquisition of new languages but can also affirm students' identities and generate a sense of belonging. Using translanguaging in the classroom is an effective way to achieve this goal. Translanguaging pedagogy can ensure that all students feel included, as it actively affirms their multiple identities and makes use of their full linguistic repertoires. The present paper attempts to synthesize recent literature on translanguaging, identify challenges related to its implementation in educational contexts and make suggestions for future directions.

Keywords: translanguaging, translanguaging pedagogy, bilingualism, multilingualism

Introduction

In a world characterized by globalization and people movement, now more than ever, classes of students from diverse cultural, ethnic and linguistic backgrounds are formed, and teachers are required to change their methods so as to create safe and inclusive teaching environments for all. Thus, the issue of fostering linguistic diversity in multilingual classrooms has been central in the field of language education. Students' first languages have often been seen as a hindrance for the learning of a target language which, in most cases, is also associated with higher prestige (Cummins *et al*., 2005). Indeed, for many years, bilinguals were viewed as the sum of two monolinguals coexisting in the same person, and their language skills were evaluated against monolingual standards (Grosjean, 2010). Their languages were considered as isolated, at best, or as a problem that can impede the learning of another target language. However, recent research has shown that bilinguals actually have a very complex linguistic configuration, and that their languages interact constantly and create a unique language system (Grosjean & Li, 2013; Pavlenko, 2014). Recent research has argued for the creation of multilingual spaces, where students' languages are used *alongside* the target language (e.g., Dagenais, Day, & Toohey, 2006; García, 2009a, 2009b; García, Skutnabb-Kangas & Torres-Guzman, 2006; Hélot & de Mejia, 2008; Hélot & Young, 2006).

The benefits of bi/multilingualism, which were for years examined in terms of monolingual standards and understood through a monolingual lens, have been highlighted (Baker, 2011; Bialystok, 2001; Cummins & Swain, 1986; Grosjean, 1982, 2010; Grosjean & Li, 2013; Lamarre, 2001; Pavlenko, 2014) and pedagogical practices now tend to focus on the *whole child* (e.g., Chow & Cummins, 2003; Cummins, 2006; Cummins *et al*., 2005; Dagenais Armand, Walsh & Maraillet, 2007; García, 2007). That is, students' existing knowledge (both cultural and linguistic) is viewed as a resource that they can build on (with the help of educators) and that can facilitate the acquisition of new languages, but can also affirm students' identities and generate a sense of belonging.

Translanguaging and translanguaging pedagogy

Translanguaging in the classroom is an effective way to achieve this goal. According to Hornberger and Link (2012:240), 'translanguaging refers broadly to how bilingual students communicate and make meaning by drawing on and intermingling linguistic features from different languages'. Bi/multilingual speakers use two or more languages *alongside* each other, when translanguaging (Greese & Blackledge, 2010:103), thus drawing on all their linguistic and communicative resources to make meaning. It is essential not to confuse translanguaging with code-switching. Although both notions refer to a shift between languages, code-switching and translanguaging are epistemologically different. The former adopts an external viewpoint to languages viewing the multilinguals' languages as distinct and separate, whereas the latter views the complete isolation of languages as impossible, as they are understood to be part of an integrated system (internal viewpoint) (Canagarajah, 2011; García & Kleyn, 2017; Grosjean, 2008). It is an approach to multilingualism that rather than focusing solely on languages, draws attention to the speakers' language *practices* that help them make sense of their world.

Bi/multilinguals use fluid language practices on a daily basis in various informal contexts, but tend to suppress this effortless flow between the languages that make up their linguistic repertoire, when they find themselves in formal settings (García *et al*., 2016). Researchers now emphasize the benefits of translanguaging and argue that it should be acknowledged and promoted in both formal and informal settings (Canagarajah, 2011; García & Kleyn, 2017; García, Johnson, & Seltzer, 2016; Greese & Blackledge, 2010). In translanguaging pedagogy, translanguaging is used *strategically* as a means to use and explore the learners' full linguistic repertoires. Teachers are encouraged to leverage the learners' communicative resources, as it is argued that this can both enhance bi/multilingual teaching, and actively affirm learners' identities (García *et al*., 2016). It is an approach that views language education as co-constructive, in the sense that knowledge is perceived as the product of teachers' and students' interaction. By affirming all learners' linguistic repertoires, translanguaging also challenges social and linguistic inequalities; it is described as the 'complex discursive practice where all students' language practices work toward sustaining ways of communicating, while giving voice to new sociopolitical realities by interrogating linguistic inequalities' (Mateus, 2014:367).

Promoting translanguaging in the classroom empowers learners by affirming their identities and enhancing their self-confidence, because it places emphasis on their complete linguistic repertoires as resources for communication and learning and views them as active agents and co-constructors of knowledge in the learning process (García & Wei 2014; Greese & Blackledge, 2010). It is especially important for marginalized students, as it can balance the power relations in classrooms (Canagarajah, 2011) and can lead to a more equitable education by challenging social and linguistic inequalities (García & Kleyn, 2017). Translanguaging can be used to facilitate the learning of a target language or to ensure that all students comprehend the class content, but this is only one of its strengths. What is perhaps even more important is that translanguaging can be used to transform learners' views of bi/multilingualism and of themselves as learners and bi/multilinguals (García *et al*., 2016).

Some practices for implementing translanguaging in the classroom include the use of dual-language books (books written by learners in two or more languages), the use of language journals (journals written by learners, in which they reflect on their language learning as well as their linguistic practices), the use of apps that allow learners to record oral languages, the reading of stories written in two or more languages, and the creation of plays where dialogues are written in various languages, thus targeting multilingual audiences (Cummins *et al*., 2005; Kirsch, 2017; Kwon & Schallert, 2016). Learners can also be encouraged to take notes in their first languages while listening or reading texts in the target language, use their first languages to ask clarification for assignments or new content, or use (or create) bilingual dictionaries (Daniel & Pachero, 2016). Evidently, teachers must model translanguaging and encourage learners to follow them, and they must also devote time and effort to create a translanguaging design for instruction, by selecting appropriate activities and material, and setting relevant learning goals (Celic & Seltzer, 2012/2013; García & Wei, 2014; García & Kleyn, 2017).

Challenges

Although translanguaging pedagogy and translanguaging research are fairly recent, and evidently empirical work and theorizing is greatly needed, some limitations have been identified. Canagarajah (2011) notes that instead of focusing solely on translanguaging practices, researchers should also study how translanguaging is seen by in-groups and out-groups, and how both groups negotiate their identities in their interactions. Canagarajah (2011) also argues that translanguaging research should not be limited to multilingual interlocutors, as multilinguals draw on their communicative resources when they interact with monolinguals (or, when the common language is English, native English speakers) too, and thus an examination of translanguaging practices in contexts where there is a mix of speakers would shed light on their communicative strategies. Indeed, recent research stresses the appropriateness of translanguaging for all contexts and among all speakers, irrespective of the number of languages that make up the repertoires of these speakers. In fact, translanguaging is unique in that it has shifted the focus from languages to language practices, and in this way interlocutors are no longer viewed as monolinguals or bi/multilinguals but rather as active agents who can use all their resources creatively in order to make meaning (Celic and Seltzer, 2012/2013; García & Wei 2014; García & Kleyn, 2017).

Researchers have also highlighted the importance of studying and promoting translanguaging in writing, as some scholars and educators find it inappropriate for formal contexts and associate translanguaging solely with oral communication (Canagarajah, 2011; Celic & Seltzer, 2012/2013). Evidently, this raises the issue of assessment and its role in translanguaging pedagogy. Standardized monolingual assessments continue to be the norm, despite the fact that adaptive technologies have allowed for the creation of translanguaged assessments. In these assessments, questions are administered in various languages and students are free to use all their linguistic resources to answer them (García, 2009; Lopez, Turkan, & Guzman-Orth, 2017). This reluctance to use translanguaged assessments reveals the hesitation of both policy makers and educators and designates the need for empirical work that will further highlight the benefits of translanguaging for learners; both monolinguals and bi/multilinguals. Educators could also be hesitant to use translanguaging and create multilingual spaces in the classroom, because they feel unprepared to do so. Indeed, there is a need for more professional development and teacher training programs that will equip teachers of all disciplines –not just language teachers- to introduce their students to translanguaging and to actively challenge widely held views that perpetuate a monolingual understanding of bi/multilingualism. The CUNY-NYSIEB project (Celic & Seltzer, 2012/2013) offers pedagogical resources and guides to educators, and further guidance can be found in the works of García *et al*. (2016) and García and Kleyn (2017).

Moving Forward - Conclusions

In recent years, many research studies have been undertaken and demonstrated that translanguaging is a common practice for bi/multilinguals in informal contexts, and can be particularly beneficial when used in formal settings too. It is an ideal way for speakers to use all their linguistic and communicative practices and it can affirm their identities. Research has shown that in the last years, many schools have started to create multilingual spaces that allow students to use their full linguistic repertoires (García *et al*. 2016). They are thus starting to move away from an understanding of languages as isolated and distinct, and towards a new understanding that places emphasis on the communicative practices that allow learners to make meaning. Using two or more languages alongside each other is beneficial for learners, as it allows them to use all their resources to understand new knowledge and communicate effectively. It is stressed, however, that translanguaging is not just a scaffold; it is not just a way for learners to learn a target language more easily, but rather it is an approach to languaging that can transform the way students and educators think about languages, learning and their own identities as bi/multilinguals.

Educators need to see themselves as co-learners and acknowledge that students can construct knowledge themselves. Translanguaging pedagogy can also help educators and students resist linguistic hierarchies and challenge long-established dominant language ideologies. Educators, however, also need guidance in terms of how to use and promote translanguaging in the classroom, as well as in terms of creating a translanguaging design for instruction. They can now benefit from projects like CUNY-NYSIEB and academic works for pedagogical resources and guides, but teacher training and professional development programs are indispensable to ensure that they are fully prepared to introduce translanguaging in their teaching.

As it has been stressed, translanguaging is appropriate for all contexts of learning and not just language learning. Therefore, more research with an emphasis on translanguaging in diverse content areas, such as science (Langman, 2014) or mathematics (He, Lai, &Lin, 2017) is needed, just as it is equally important to explore the possibilities of translanguaging offered by technology (Schreiber, 2015). In addition, researchers should also explore in depth the use of translanguaging in the context of heritage language (an example being Golneshan, 2015) and its relation to the phenomena of language shift and language maintenance. Evidently, more action needs to be taken at the level of language policies that account for translanguaging and allow for the creation of multilingual spaces in schools and other institutions.

References

Baker, C. (2011). *Foundations of bilingual education and bilingualism* (5th edn). Toronto: Multilingual Matters.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy and cognition*. New York: Cambridge University Press.

Canagarajah, S. (2011). Translanguaging in the classroom: Emerging issues for research and pedagogy. *Applied Linguistics Review*, *2*, 1-28.

Celic, C., & Seltzer K. (2012/2013). *Translanguaging: A CUNY-NYSIEB guide for educators*. New York: CUNY-NYSIEB, The Graduate Center, The City University of New York. Retrieved on October 17 2017 from: http://www.nysieb.ws.gc.cuny.edu/files/2013/03/Translanguaging-Guide-March-2013.pdf

Chow, P., & Cummins, J. (2003). Valuing multilingual and multicultural approaches to learning. In S. R. Schecter & J. Cummins (eds.), *Multilingual education in practice: Using diversity as a resource* (pp. 32-55). Portsmouth, NH: Heinemann.

Cummins, J. (2006). Identity texts: The imaginative construction of self through multiliteracies pedagogy. In O. García, T. Skutnabb-Kangas, & M. E. Torres-Guzman (eds.), *Imagining multilingual schools: Languages in education and globalization* (pp. 51-68). Clevedon, UK: Multilingual Matters.

Cummins, J., Bismilla, V., Chow, P., Cohen, S., Giampapa, F., Leoni, L., Sandhu, P., & Sastri, P. (2005). Affirming Identity in Multilingual Classrooms. *Educational Leadership*, *63*(1), 38-43.

Cummins, J., & Swain, M. (1986). *Bilingualism in education: Aspects of theory, research and practice*. London: Longman.

Dagenais, D., Armand, F., Walsh, N., & Maraillet, E. (2007). L’ éveil aux langues et la co-construction de connaissances sur la diversité linguistique. *Canadian Journal of Applied Linguistics*, *10*(2), 197-219.

Dagenais, D., Day, E., & Toohey, K. (2006). A multilingual child’s literacy practices and contrasting identities in the figured worlds of French immersion classrooms. *International Journal of Bilingual Education and Bilingualism*, *9*(2), 205-218.

Daniel, S. M., & Pacheco, M. B. (2016). Translanguaging practices and perspectives of four multilingual teens. *Journal of Adolescent and Adult Literacy*, *59*(6), 653-663.

García, O. (2009a). *Bilingual education in the 21st century: A global perspective*. Malden, MA: Wiley-Blackwell.

García, O. (2009b). Education, multilingualism and translanguaging in the 21st century. In A. Mohanty, M. Panda, R. Phillipson, & T. Skutnabb-Kangas (eds.), *Multilingual education for social justice: Globalising the local* (pp. 140-158). New Delhi: Orient Blackswan.

García, O. (2007). Foreword: Intervening discourses, representations and conceptualizations of language. In S. Makoni & A. Pennycook (eds.), *Disinventing and reconstituting languages* (pp. 11-15). Clevedon, UK: Multilingual Matters.

García, O., & Kleyn, T. (eds.) (2017). *Translanguaging with multilingual students: Learning from classroom moments*. New York: Routledge.

García, O., Johnson, S., & Seltzer, K. (2016). *The translanguaging classroom. Leveraging student bilingualism for learning*. Philadelphia: Caslon.

García, O., Skutnabb-Kangas, T., & Torres-Guzman, M. E. (2006). *Imagining multilingual schools: Languages in education and globalization*. Clevedon, UK: Multilingual Matters.

Golneshan, Z. (2016). *Exploring translanguaging practice in creating digital identity texts in a Farsi heritage language class*. PhD dissertation retrieved on October 20 2017 from http://theses.ucalgary.ca /bitstream/ 11023/3521/1/ucalgary_2016_golneshan_zahra.pdf

Greese, A., & Blackledge, A. (2010). Translanguaging in the bilingual classroom: A pedagogy for learning and teaching? *The Modern Language Journal*, *94*(1), 103-115.

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press.

Grosjean, F. (2008). *Studying bilinguals*. Oxford, UK: Oxford University Press.

Grosjean, F. (2010). *Bilingual: Life and reality*. Cambridge, MA: Harvard University Press.

Grosjean, F., & Li, P. (2013). *The psycholinguistics of bilingualism*. New York, NY: John Wiley.

He, P., Lai, H., & Lin, A. (2017). Translanguaging in a multimodal mathematics presentation. In C. M. Mazak & K. S. Carroll (eds.), *Translanguaging in higher education: Beyond monolingual ideologies* (pp. 91-120). Bristol, UK: Multilingual Matters.

Hélot, C., & de Mejía, A. M. (eds.). (2008). *Forging multilingual spaces: Integrated perspectives on majority and minority bilingual education*. Clevedon, UK: Multilingual Matters.

Hélot, C., & Young, A. (2006). Imagining multilingual education in France: A language and cultural awareness project at primary level. In O. García, T. Skutnabb-Kangas, & M. E. Torres-Guzman (eds.), *Imagining multilingual schools: Languages in education and globalization* (pp. 69-90). Clevedon, UK: Multilingual Matters.

Hornberger N. H., & Link, H. (2012). Translanguaging in today's classrooms: A biliteracy lens. *Theory Into Practice*, *51*(4), 239-247.

Kirsch, C. (2017). Translanguaging practices during storytelling with the app iTEO in preschools. *Translation and Translanguaging in Multilingual Contexts*, *3*(2), 145-166.

Kwon, H. J., & Schallert, D. L. (2016). Understanding translanguaging practices through a biliteracy continua framework: Adult biliterates reading academic texts in their two languages. *Bilingual Research Journal*, *39*(2), 138-151.

Lamarre, P. (2001). Plurilingualism and pluriculturalism: An approach from the Canadian perspective. *Kolor: Journal on Moving Communities*, *1*(1), 33-45.

Langman, J. (2014). Translanguaging, identity, and learning: Science teachers as engaged language planners. *Language Policy*, *13*(2), 183-200.

Lopez, A. A., Turkan, S., & Guzman-Orth, D. (2017). Conceptualizing the use of translanguaging in initial content assessments for newly arrived emergent bilingual students. *ETS Research Report Series*. Retrieved on October 27 from: http://onlinelibrary.wiley.com/doi/10.1002/ets2.12140/pdf

Mateus, S. G. (2014). Translanguaging: Language, bilingualism, and education. *Bilingual Research Journal*, *37*(3), 366-369.

Pavlenko, A. (2014). *The Bilingual mind and what it tells us about language and thought*. Cambridge, MA: Cambridge University Press.

Schreiber, B. R. (2015). ‘I am what I am’: Multilingual identity and digital translanguaging. *Language Learning and Technology*, *19*(3), 69-87.

To cite this article: Tisizi, A. (2017). Translanguaging: A synthesis of recent literature and suggestions for future directions of research and action. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 271-275). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Approaches to the assessment and treatment of speech sound disorders in children among language therapists in Israel

Gila Tubul-Lavy[1], Ariela Nachamani[2], Limor Adi-Bensaid[1]
gilatu@gmail.com, ariela50@bezeqint.net, limorbensaid@gmail.com

[1]Ono Academic College, [2]Hadassah Academic College

Abstract. Speech sound disorder (SSD) includes speech impairment due to motor and/or cognitive-phonological impairment. Because of the heterogeneity of the disability, methods of assessment and treatment are varied, and speech language therapists (SLTs) find it difficult to decide which method to choose. Web questionnaires are one way to examine how SLTs assess and treat children with SSD and are used in various countries (Joffe & Pring, 2008; Oliveira, Lousada, & Jesus, 2015). The goal of the present study was to examine the attitude of SLTs in Israel in the field of SSD and to describe the most common assessment and treatment approaches used with these children. The data were obtained from a web questionnaire completed by 93 SLTs comprising 20 questions relating to their fields of work, the patients' ages, the treatment methods they currently use, and their thoughts on the treatment of SSD. The data collected enable a contemporary and up-to-date mapping of this field among Israeli SLTs. The findings show that the assessment methods preferred by Israeli SLTs include the assessment of word pronunciation, syllable imitation, and auditory discrimination, and the analysis of spontaneous speech and oro-motor abilities. These results are compared to results of previous questionnaires used in various countries.

Keywords: speech sound disorders, questionnaire, motor impairment, phonological impairment, childhood apraxia of speech

Introduction

Speech sound disorder (SSD) is defined as speech impairment resulting from motor and/or cognitive-phonological difficulty that interferes with verbal communication, social participation, academic performance or occupational performance due to low speech comprehension (APA, 2013).

A high percentage of children with SSD are referred for speech therapy: 10-15 percent of kindergarten children and six percent of school children (McLeod & Harrison, 2009). At school, comorbidities of 25-30 percent were found for reading impairment and SSD, rising to 66 percent in children with both SSD and language disorders (Anthony, Aghara, Dunkelberger, Anthony, Williams, & Zhang, 2011).

The methods used to assess and treat SSD may be based on motor or phonological principles. However, the variability between the characteristics, severity, and etiology of the disorder sometimes raises doubts regarding effective treatment approaches and goals. The goal of the traditional articulation approach (Van Riper, 1947) is to assess articulation of individual phonemes and to teach the child how to improve the perception and production of phonemes. However, phonological assessment focuses on searching error patterns and changing the child's phonological/language system (Dodd & Bradford, 2000). Thus, instead of changing each phoneme production in phonemic therapy, the aim of phonological therapy is to work with groups of sounds in words (McLeod & Baker, 2016). Another type of SSD is childhood apraxia of speech (CAS) which relates to the difficulty in planning and carrying out deliberate movements (Crary, 1993; Steinman, Mostofsky, & Denckla, 2010). In most cases, CAS includes both motoric and phonological impairment.

Several studies have used questionnaires to examine the different assessment and treatment methods of SSD employed by SLTs from different countries (Joffe & Pring, 2008 in the United Kingdom; Oliveira *et al*., 2015 in Portugal; Pascoe, Maphalala, Ebrahim, Hime, Mdladla, Mohamed, & Skinner, 2010 in South Africa).These questionnaires were adapted to suit each country's specific treatment culture.

Most of the SLTs in UK use a formal assessment (The South Tyneside Assessment of Phonology, Armstrong & Ainley, 1992) and they reported confidence in choosing treatment methods which include auditory discrimination, minimal contrast therapy, and phonological impairment, often used in combination. Most of them involve parents in assessment and therapy (parent-based work). In choosing the intervention plan, SLTs seem to be more influenced by children's language and cognitive abilities and the motivation of parents than by the nature of the impairment (Joffe & Pring, 2008).

In South Africa, almost 90% of the SLT's use an integration of informal with formal assessment. They use a variety of interventions such as auditory discrimination and phonological awareness, often in combination, and are based on a child's profile of difficulties. Because most of the children in South Africa are bilingual, half of the SLTs felt unsure about the selection of assessments and intervention for these children (Pascoe *et al*., 2010).

In Portugal, only intervention was examined. Integration of expressive phonological tasks, phonological awareness, listening and discrimination activities was found to be an effective method training for SSD (Oliveira *et al*., 2015). In Israel, there is no formal test for assessing SSD. Carter, Lees, Murira, Gona, Neville, & Newton (2004) emphasize the need to develop cross-cultural assessments of speech and language for children, but to this day every country and its culture use a different assessment.

Aim and method

The present study examined the views of Israeli SLTs on SSDs of among children and the common approaches they use in assessment and treatment.

An online questionnaire was developed by the authors, three-experienced SLTs who have worked with SSD for over 30 years. The questionnaire was based on previous questionnaires developed in British-English (Joffe & Pring, 2008), and Portuguese (Oliveira *et al*., 2015) and was adapted to Israeli culture.

This questionnaire included 20 closed and open-ended questions. Four questions addressed the characteristics of the work of SLTs (e.g., 'How many years have you been working as a clinician?'); six addressed methods of assessment and distinction among SSD (e.g., 'Rate the criteria from the table below that help to distinguish between phonological disorders and CAS'); eight related to intervention (e.g., 'How confident are you about choosing the right treatment when planning a treatment plan for a patient with the following disorders: phonetic disorders, phonological disorder, CAS?'); while the final two questions concerned evidence-based practice (e.g., 'How much do you agree with the following statement: In my opinion, there is evidence that shows the effectiveness of intervention in the treatment of phonological disorders').

The questionnaire was distributed on the Israeli SLT's website and through the Ministry of Health, HMOs, schools, and kindergartens, as well as private clinics. It was completed by 93 SLTs who work with children with SSD.

Results

The first goal was to review the assessment methods used by Israeli SLTs for children with SSD. Figure 1 presents the average of preferred assessment method by Israeli SLTs. (Range between 0-5: 5-always, 1-never, 0-not familiar with this method).

The second goal of the study was to review accepted treatment methods. Figure 2 presents the mean value of preferred treatment method among Israeli SLTs (on a scale of 0-5: 5-always, 1-never, 0-not familiar with this method).

A further question addressed the self-confidence of SLTs in choosing the right treatment plan for three SSD types: articulation disorder, phonological disorder, and CAS. Figure 3 presents the answers to this question.

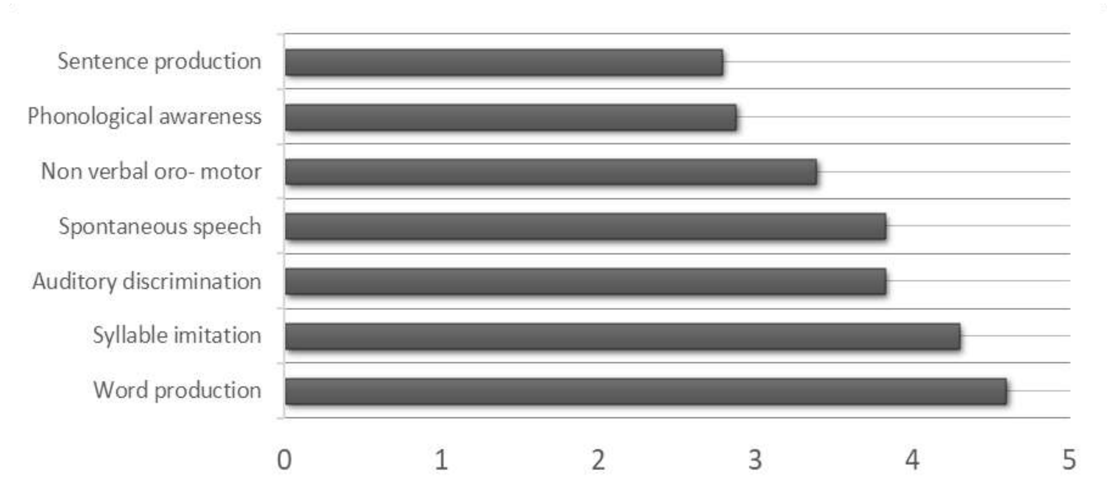


Figure 1: Mean value for preferred assessment procedure

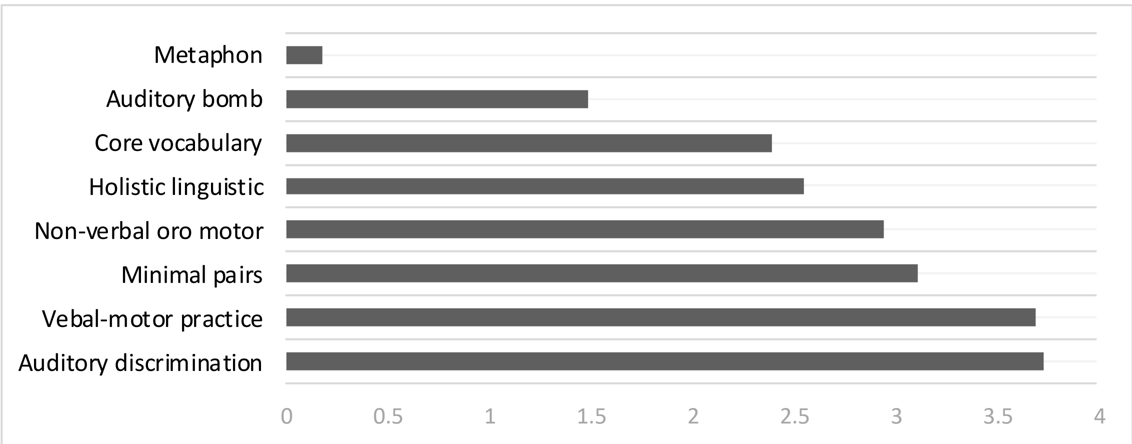


Figure 2: Mean value for preferred treatment procedure

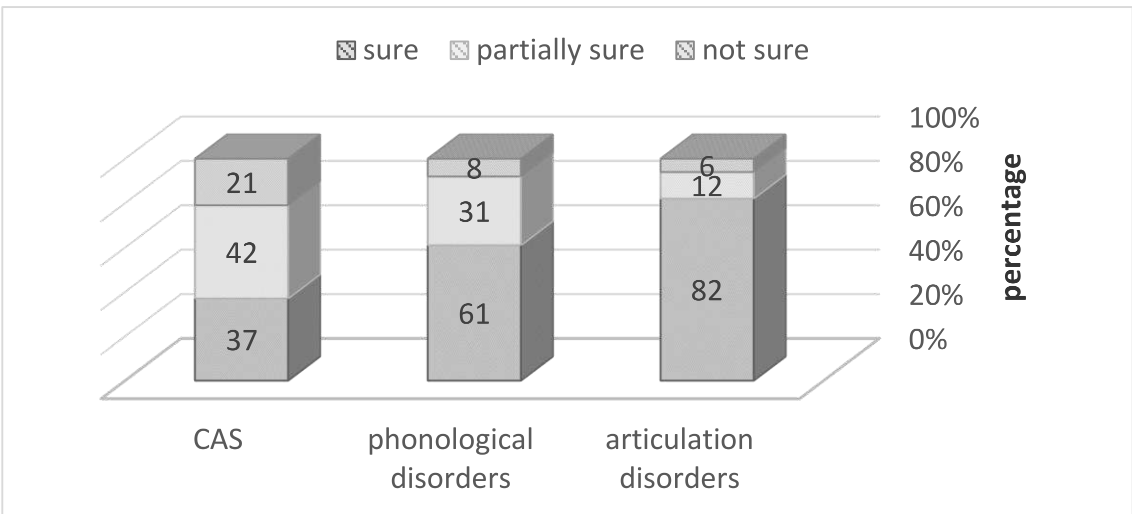


Figure 3: Confidence in choosing an appropriate treatment plan for SSD (percentage)

Discussion

Although SSD is a common disorder treated by SLTs, there is still lack of agreement concerning the approaches that may be used in assessment and intervention with children with SSD (Baker & McLeod, 2011).

Assessment

Most SLTs in the UK tend to use a quick and easy screening tool to assess speech (*Assessment of South Tyneside Assessment of Phonology*, Armstrong & Ainley, 1992). In South Africa, ten different formal assessments are used to assess children's speech, most commonly the Goldman-Fristoe articulation test (Goldman & Fristoe, 1986). In Israel there is no formal test for assessing SSD, and this is almost certainly one of the reasons why each SLT employs different informal methods.

Intervention

Some of our findings are partially compatible with previous questionnaires in other countries. In Israel, as in England (Joffe & Pring, 2008), Portugal (Oliveira *et al*., 2015) and South Africa (Pascoe *et al*., 2010), auditory discrimination and minimal pairs were found to be the most common treatment approaches. However, some differences were found between the countries: Core vocabulary approach was used frequently in South Africa for bilingual children. In the UK and Israel, it was used rarely but for monolingual children. The SLTs in Israel used parent-based work less than their peers in other countries and they utilized more non-speech oro-motor. In all questionnaires, and most clinicians stated that they use more than one approach.

Confidence in choosing treatment methods appears to depend on the severity of the impairment: Israeli SLTs seem to be more confident in choosing treatment methods for articulation rather than phonological disorders, and 63 percent of them were uncertain about the appropriate treatment methods for CAS.

SLTs in the UK who used structured diagnosis reported greater confidence in choosing a treatment method than those using general assessments. Since there is no structured diagnosis for SSD in Hebrew, this may explain the lack of confidence among the Israeli SLTs, particularly in assessing severe impairments.

Conclusion

Our findings enable a comparison between the clinical experiences of SLTs from different countries, as well as between different studies. This study led to a decision to attempt to enhance Israeli SLTs' awareness of the assessments and treatments available for this population.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM - 5®)*. American Psychiatric Pub.

Anthony, J. L., Aghara, R. G., Dunkelberger, M. J., Anthony, T. I., Williams, J. M., & Zhang, Z. (2011). What factors place children with speech sound disorders at risk for reading problems? *American Journal of Speech-Language Pathology*, *20*(2), 146-160.

Armstrong, S., & Ainley, M. (1992). *South Tyneside assessment of phonology*. Ponteland: STASS Publications.

Baker, E., & McLeod, S. (2011). Evidence-based practice for children with speech sound disorders: Part 1 narrative review. *Language, Speech, and Hearing Services in Schools*, *42*(2), 102-139.

Carter, J., Lees, J., Murira, G., Gona, J., Neville, B., & Newton, C. (2004). Issues in the development of cross-cultural assessments of speech and language for children. *International Journal of Language and Communication Disorders*, *40*(4), 385-401.

Crary, M. A. (1993). *Developmental motor speech disorders*. San Diego, CA: Singular Publishing Group.

Dodd, B., & Bradford, B. D. A. (2000). A comparison of three therapy methods for children with different types of developmental phonological disorder. *International Journal of Language and Communication Disorders*, *35*(2), 189-209.

Goldman, R., & Fristoe, M. (1986). *Goldman Fristoe test of articulation*. Circle Pines, MN: American Guidance Service.

Joffe, V., & Pring, T. (2008). Children with phonological problems: A survey of clinical practice. *International Journal of Language and Communication Disorders*, *43*(2), 154-164.

McLeod, S., & Harrison, L. J. (2009). Epidemiology of speech and language impairment in a nationally representative sample of 4- to 5-year-old children. *Journal of Speech, Language, and Hearing Research*, *52*(5), 1213-1229.

McLeod, S., & Baker, E. (2016). *Children's speech: An evidence-based approach to assessment and intervention*. Boston, MA: Pearson Education.

Oliveira, C., Lousada, M., & Jesus, L. M. (2015). The clinical practice of speech and language therapists with children with phonologically based speech sound disorders. *Child Language Teaching and Therapy*, *31*(2), 173-194.

Pascoe, M., Maphalala, Z., Ebrahim, A., Hime, D., Mdladla, B., Mohamed, N., & Skinner, M. (2010). Children with speech difficulties: A survey of clinical practice in the Western Cape. *The South African Journal of Communication Disorders*, *57*(1), 66-75.

Steinman, K. J., Mostofsky, S. H., & Denckla, M. B. (2010). Toward a narrower, more pragmatic view of developmental dyspraxia. *Journal of Child Neurology*, *25*(1), 71-81.

Van Riper, C. (1947). *Speech correction: Its principles and methods*. Engelwood Cliffs, NJ: Prentice-Hall.

To cite this article: Tubul-Lavy, G., Nachamani, A., & Adi-Bensaid, L. (2017). Approaches to the assessment and treatment of speech sound disorders in children among language therapists in Israel. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 276-280). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Theory of Mind and comprehension of factivity in specific language impairment: Perspectives from an intervention study

Maria Voulgaraki, Stavroula Stavrakaki
marigakivou@gmail.com, svoula@itl.auth.gr

Aristotle University of Thessaloniki

Abstract. This paper explores the relation of theory of mind (ToM) abilities to the reception of factivity by assessing the comprehension of factiviy and ToM skills in children with specific language impairment (SLI) and evaluating the effects of linguistic intervention on the performance of children with SLI on ToM abilities and factivity comprehension. The participants with SLI were tested at three different times: before intervention (t1), immediately after intervention (t2), as well as 8 months later (t3). Our results indicated that linguistic therapy restricted to factivity has beneficial effects on ToM performance, in addition to the reception of factivity per se. While these effects decreased over time as shown by re-assessment at t3, they did not completely disappear. We conclude that semantic therapy can be generalized in different domains and have positive impact on them.

Keywords: specific language impairment, theory of mind, factivity, linguistic intervention

Introduction

Specific language impairment (SLI) is a disorder of language acquisition in the presence of typical non-verbal IQ abilities. While deficits in language and especially grammar constitute a 'hallmark' for children with SLI, other cognitive abilities, particularly verbal short-term memory, appear to be significantly affected, as well. In addition, SLI often co-occurs with symptoms of other developmental disorders, for example, developmental dyslexia or autism (Rice, 2016). Current research has paid particular attention to the interaction of linguistic and cognitive abilities in SLI. Furthermore, it has explored whether core symptoms of other neurodevelopmental disorders appear in SLI (Leonard, 1998/2014). Theory of mind, the ability to attribute independent mental states to ourselves and others (Baron-Cohen, Leslie, & Frith, 1985), has been proved a fertile domain to investigate the aforementioned issues. This is so for the following reasons.

First, it has been extensively argued that ToM is closely related to linguistic development, as significant correlations between standardized language measures and performance on ToM tasks have been reported for preschoolers (see Cutting & Dunn, 1999, among others). Particular attention has been paid to the interrelation of ToM and comprehension of embedded syntax (complement clauses) (de Villiers & de Villiers, 2000). Furthermore, it has been suggested that the acquisition of semantic properties of mental state verbs are related to the development of ToM (de Villiers & Pyers, 1997). Especially, the distinction between factive (e.g., *know*) and non-factive (e.g., *think*) predicates is closely related to the ability to understand thoughts and behavior of ourselves and others (Schulz, 2003). Notably, factive and non-factive predicates differ crucially in their semantic properties: While the truth of the complement embedded clause is presupposed in factive predicates (see example 1), this is not the case for non-factive predicates (see example 2) (Kiparsky & Kiparsky, 1970).

(1) I *know* John is in his room.

(2) I *think* John is in his room.

Second, ToM has been particularly deficient in autism (Baron-Cohen, Leslie & Frith, 1985) and considered as the cause of the communication impairment in this disorder. It is, thus, of particular interest to investigate whether ToM is affected in another neurodevelopmental disorder, in particular, SLI. Research findings indicate that children with SLI have significantly lower performance than typically developing peers on ToM, mainly on false-belief tasks (Jester & Johnson, 2016). Most researchers argue that poor performance on ToM is largely due to linguistic deficits (Farrant, 2015) and suggest that ToM and grammar (Loukusa, Mäkinen, Kuusikko-Gauffin, Ebeling, & Moilanen, 2014) or vocabulary (Farrar, Johnson, Tompkins, Easters, Zilisi-Medus, & Benigno, 2009) are interrelated.

The present study addresses the relation of ToM abilities to linguistic (semantic) abilities from an innovative angle, as it examines whether semantic intervention focusing on the reception of factivity has beneficial effects on the reception of factivity per se as well as on ToM abilities. Furthermore, it explores whether long-term beneficial effects follow a specific focused intervention by testing the participants with SLI before intervention (*time (t)1*), immediately after intervention (*t2*) and 8 months after intervention (*t3*). Based on previous research showing relation between ToM and linguistic skills, we hypothesize that training in the reception of factivity should positively affect the performance of children with SLI on ToM tasks in addition to factivity per se. As only a few studies have investigated whether the intervention effects can be long-term (for a review: Law Garrett & Nye, 2010), it is crucial to see the extent to which the children of this study benefited from the treatment they received during intervention. Given that assessment at t3 was performed after an almost three-month summer holiday and children with SLI did not receive therapy over this period, we would like to see whether any beneficial intervention outcomes could remain in the absence of any intervention stimulus.

Method

Participants

Eight children, aged 7-11 years old, diagnosed with SLI by multidisciplinary teams in Thessaloniki, participated in our study. All children attended additional special classes in the mainstream school, while only two of them received speech and language therapy services at the time of testing. None of these two children received speech and language therapy focusing on the semantics of factive and non-factive predicates or on ToM by their therapists.

Table 1. Participant information and clinical evaluation results

Participants	Age	Gender	Verbal IQ (Raven)	Non-verbal IQ (Raven)	Action picture test (grammar)	Action picture test (information)	Action picture Test (total)	Sentence repetition (DVIQ test)
G.	7;1	M	85	95	32	37	69	41
T.	7;9	M	85	115	29	38	67	37
A.	8;3	M	85	95	34	42	76	41
M.	8;4	M	80	85	33	43	76	41
P.	8;7	M	65	85	37	48	85	40
K.	9;1	M	85	95	34	43	77	43
E.	11;2	F	70	100	39	43	82	44
S.	11;3	F	70	95	35	44	79	45

M: male, F: female

Pre-experimental assessment was performed by employing the following materials: (i) the Greek version of the Raven test (Sideridis, Antoniou, Mouzaki, & Simos, 2015) to assess their verbal and non-verbal IQ (ii) the sentence repetition (SER) task (subpart of the preschool version of the Diagnostic Verbal IQ (DVIQ) Test) (Stavrakaki & Tsimpli, 2000) to assess verbal short term memory and (iii) the Greek version of the Action Picture Test (APT^{GR}) (Vogindroukas, Protopappas, & Stavrakaki, 2009), to assess pragmatic and grammar abilities. All children were individually tested at their school by the first author of this paper (MV). Individual data for SLI participants including gender, age and performance on clinical measures are shown in Table 1.

As Table 1 shows, all participants with SLI showed normal non-verbal IQ while their verbal performance on the relevant section of RAVEN^{GR} was apparently below their non-verbal IQ abilities. Additional testing of their linguistic abilities by the means of APT^{GR} confirmed their difficulties in the verbal domain (Vogindroukas *et al*., 2009). Specifically, the participants with SLI used smaller vocabularies than those expected given their CA while they avoided producing embedded clauses and showed impairment in weak pronouns (clitics). Furthermore, all children with SLI, except child S, showed difficulties in SER.

Pre-therapeutic experimental evaluation

We employed experimental tasks to assess ToM abilities and factivity reception. With respect to the former, we tested false belief understanding with Sally-Ann (Wimmer & Perner, 1983) and Smarties test (Perner, Frith, Leslie, & Leekam, 1989). With respect to the latter, we developed a tool, in particular, the factivity reception task (FART), based on the task developed by Moore, Bryant, and Furrow (1989). These researchers evaluated understanding of factivity by assessing whether participants were able to distinguish the verbs *know* and *think*. The assessment procedure was as follows. The examiner was hiding an object (e.g., a sweet) in one of two locations (e.g., a blue and a red box) and the child had to find where the sweet was, based on the statements made by two dolls. For example, one doll said: ‘I know it is in the blue box’, while the other said: ‘I think it is in the red box’. FART consisted of ten experimental stories as well as five distractor stories inserted at random order amongst the main experimental stories (see Voulgaraki, 2017 for details). To validate FART, as it was not previously employed for Greek, we tested 17 typically developing children (TDC), monolingual speakers of Greek aged 7-9.5 years. Their performance on FART is shown in Table 2, while performance of the SLI participants on the experimental tasks in t1 is presented in table 3.

Table 2. TDC’s performance on FART

Number of participants	CA Mean (SD)	FART score Mean (SD)
17	8;1 (0.72)	8.7 (2.05)

Table 3. The performance of children with SLI on Theory of Mind[1] and FART[2] in t1, t2, and t3

Participants	FART (t1)	FART (t2)	FART (t3)	ToM (t1)	ToM (t2)	ToM (t3)
G.	3	10	5	0	1	2
T.	2	10	7	1	2	1
M.	2	10	4	0	0	0
P.	3	10	9	1	2	2
K.	2	10	10	1	2	2
S[3]	4	9	-	0	2	-

[1]: maximum score: 2

[2]: maximum score: 10

[3]: The participant S was not included in t3, as it was not possible to find her.

At t1, six out of eight students showed significant deficits in ToM and in factivity reception, which are not expected given their chronological age; ToM as well as a distinction between factive and non-factive predicates are expected to be acquired around the ages of 4-5 (Happé, 1995; Moore *et al*., 1989). Children who performed at ceiling on both tasks (child E) or showed high level of performance on one out of the two experimental tasks (child A) were excluded from intervention.

Intervention

We developed an intervention programme aimed at distinguishing factive from non-factive verbs (Voulgaraki, 2017, for details) based on previous intervention methodologies, namely, on intervention by *color semantics* (Bryan, 1997), *shape coding* (Ebbels, 2007) and *thought bubbles* which visually present thoughts (Wellman & Peterson, 2013). Specifically, factive and non-factive verbs were presented with different shapes and colors to children with SLI. The same verbs were employed in intervention as in the experimental setting. However, while scenarios with dolls were employed for experimental assessment, stories supported by visualised thought bubbles were used in the intervention setting. In particular, the child was firstly presented with these stories and then trained to distinguish factive from non-factive predicates by picking the appropriate verb up after the presentation of a scenario. In Appendix I, a figure of intervention material is presented. The participant was immediately informed of the correctness or otherwise of her/his response and presented with the correct answer. In addition, each child was encouraged to develop a story by selecting a bubble thought and the appropriate verb. Corrective feedback was also provided in this case. All children received therapy individually by MV at their school for 8 times over a period of one and a half month. Each session lasted approximately 30 minutes.

Re-assessment

All participants were re-assessed on the same experimental tasks as in pre-therapeutic assessment immediately after intervention, as well as 8 months after intervention to investigate short- and long-term intervention effects, respectively. Assessment at t3 was performed after the summer holiday and, as children did not receive therapy over this period (extended to almost three months), we would like to see whether any beneficial therapy outcomes remain. In Table 3, the performance of children with SLI on FART and ToM tasks is shown before intervention (t1), immediately after intervention (t2), and after 8 months (t3).

Table 3 indicates that all children with SLI showed significant improvement in factivity reception at t2. Non-parametric statistics (within group analysis) confirmed these results (Z = 2.232, p = 0.026). This was also the case for ToM tasks (t1 vs. t2: Z = 2.121, p = 0.034). At t3, the participants' performance dropped in the FART and ToM tasks compared to their performance at t2. However, drop in FART at t3 compared to t2 just failed to be significant (Z=1,826, p=0,068), while it was not significant at all in ToM tasks (Z=0,0 p=1). Notably, the children with SLI still performed significantly better at t3 than at t1 in FART (Z=2.03, p=0,042), though this was not the case for ToM tasks (t1 vs. t3:Z=1. 63,p=0.102).

Discussion

This study investigated whether children with SLI showed difficulties in the domain of ToM and factivity reception. It also explored whether reception of factivity related to ToM from an innovative perspective by performing intervention in the domain of factivity. Furthermore, we examined the extent to which any beneficial intervention effects remain over time by performing a follow-up re-assessment eight months after the intervention programme. Our initially testing confirmed previous findings showing difficulties in ToM and factivity reception in SLI (Farrant, 2015; Loukusa *et al*., 2014). Specifically, most (but not all) of our participants showed such difficulties. The intervention applied showed immediate beneficial effects for the participants with SLI whose group performance was significantly improved in both experimental tasks. Notably, one participant (M) showed no improvement in ToM at all at t2 (after intervention; recall that the non-verbal IQ of this participant was not high, although normal). It might be interesting to see whether performance on ToM tasks can be related to non-verbal abilities and, if so, to what extent (Bock, Gallaway, & Hund, 2015). Apparently, this question cannot be answered from the data of this study.

Notably, the results showed parallel recovery in both domains confirming that intervention effects in one domain (factivity) can be successfully transferred in a related cognitive domain (ToM). In this respect, our study provides further support to the view that factivity is interrelated to ToM (Miller, 2001). Our findings for the effectiveness of a linguistic orientated intervention method for the development of ToM abilities are in agreement with previous studies, which show that exposure to linguistic input can positively affect ToM. In particular, including dialogue for mental states in the intervention procedure, teaching conversational skills, or reading stories rich with mental terms have been shown to be effective and, thus, applicable in education (Bianco & Lecce, 2016; Rakoczy, Tomasello, & Striano, 2010). While we cannot exclude the possibility that better performance at t3 is due to maturation effects, as pointed out by an anonymous reviewer, apparently we cannot assess the maturation effects per se. Language development proceeds very slowly in children with SLI and, thus, it is not in accordance with general maturation (Leonard, 1998/2014). Re-assessment after eight months showed that, while participants performed lower than in the first post-intervention assessment (t2), they still performed higher than in the pre-intervention assessment (t1). Crucially, only a few studies have investigated the long terms effects of speech and language intervention in specific language impairment (Law *et al*., 2010). In addition, a recent metanalysis of 32 studies by Hofmann, Doan, Sprung, Wilson, Ebesutani, Andrews, Curtiss, and Harris (2016) showed that intervention effectiveness in the domain of ToM is highly dependent on the duration of intervention. Specifically, the longer the intervention lasts, the better outcomes it will have.

To summarize, our study provided evidence that focused linguistic intervention in the domain of factivity can improve the performance of children with SLI on ToM in addition to factivity per se. Interestingly, these beneficial results did not disappear at the end of intervention but were still present in a re-assessment performed after eight months.

References

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a ‘theory of mind’? *Cognition*, *21*, 37-46.

Bianco, F., & Lecce, S. (2016). Translating child development research into practice: Can teachers foster children's Theory of Mind in primary school? *British Journal of Educational Psychology*, *85*(4), 592-605.

Bock, A. M., Gallaway, K. C., & Hund, A. M. (2015). Specifying links between executive functioning and Theory of Mind during middle childhood: Cognitive flexibility predicts social understanding. *Journal of Cognition and Development*, *16*(3), 509-521.

Bryan, A. (1997). Colourful semantics. In S. Chiat, J. Law & J. Marshall (eds.), *Language disorders in children and adults: Psycholinguistic approaches to therapy* (pp. 143-161). London: Whurr.

Cutting, A. L., & Dunn, J. (1999). Theory of Mind, emotion understanding, language, and family background: individual differences and interrelations. *Child Development*, *70*(4), 853-65.

de Villiers, J. G., & de Villiers, P. A. (2000). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs. (eds.), *Children’s reasoning and the mind* (pp. 189-226). Hove, UK: Psychology Press.

de Villiers, J., & Pyers, J. (1997). *Complementing cognition: The relationship between language and Theory of Mind.* In *Proceedings of the 21st annual Boston University conference on language development*. Somerville, MA: Cascadilla Press.

Ebbels, S. (2007). Teaching grammar to school-aged children with specific language impairment using Shape Coding. *Child Language Teaching and Therapy*, *23*(1), 67-93.

Farrant, B. (2015). Specific language impairment and perspective taking: Delayed development of Theory of Mind, visual and emotional perspective taking. *Journal of Childhood and Developmental Disorders*, 1-8.

Farrar, M. J., Johnson, B., Tompkins, V., Easters, M., Zilisi-Medus, A., & Benigno, J. P. (2009). Language and theory of mind in preschool children with specific language impairment. *Journal of Communication Disorders*, *42*(6), 428-441.

Happé, F. (1995). *Autism: An introduction to psychological theory*. Cambridge, MA: Harvard University Press.

Hofmann, S., Doan, S. N., Sprung, M., Wilson, A., Ebesutani, C., Andrews, L., Curtiss, J., & Harris, P. L. (2016). Training children's theory-of-mind: A meta-analysis of controlled studies. *Cognition*, *150*, 200-212.

Jester, M., & Johnson, C. J. (2016). Differences in theory of mind and pretend play associations in children with and without Specific Language Impairment. *Infant and Child Development*, *25*(1), 24-42.

Kiparsky, P., &. Kiparsky, C. (1968). Fact. In M. Bierwisch & K. E. Heidolph (eds.), *Progress in linguistics* (pp. 143-173). The Hague: Mouton.

Law, J., Garrett, Z., & Nye, C. (2010). Speech and language therapy interventions for children with primary speech and language delay or disorder. *The Cochrane Database of Systematic Review*, *5*, CD004110.

Leonard, L. (1998/2014). *Children with specific language impairment.* Cambridge, MA: The MIT Press.

Loukusa S., Mäkinen, L., Kuusikko-Gauffin, S., Ebeling, H., & Moilanen, I. (2014). Theory of mind and emotion recognition skills in children with specific language impairment, autism spectrum disorder and typical development: group differences and connection to knowledge of grammatical morphology, word-finding abilities and verbal working memory. *International Journal of Language and Communication Disorders*, *49*(4), 498-507.

Miller, C. A. (2001). False belief understanding in children with Specific Language Impairment. *Journal of Communication Disorders*, *34*, 73-86.

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *Child Development*, *60*, 167-171.

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R., (1989). Exploration of the autistic child's theory of mind: Knowledge, belief, and communication. *Child Development, 60,* 689-700.

Rakoczy, H., Tomasello, T., & Striano, T. (2010). The role of experience and discourse in children’s developing understanding of pretend play actions. *British Journal of Developmental Psychology*, *24*, 302-335.

Rice, M. L. (2016). Specific language impairment, nonverbal IQ, attention-deficit/hyperactivity disorder, autism spectrum disorder, cochlear implants, bilingualism, and dialectal variants: Defining the boundaries, clarifying clinical conditions, and sorting out causes. *Journal of Speech, Language, and Hearing Research*, *59*(1), 122-132.

Schulz, P. (2003). *Factivity: Its nature and aquisition*. Tubingen: Niemeyer. https://www.uni-frankfurt.de/52179962/Factivity-Its-Nature-and-Acquisition.pdf Accessed on the 17th of October, 2017.

Sideridis, G., Antoniou, F., Mouzaki, A., & Simos, P. (2015). *RAVEN'S Progressive Matrices and vocabulary test.* Athens: Motivo (in Greek).

Stavrakaki, S., &Tsimpli, I. M. (2000). Diagnostic verbal IQ test for Greek preschool and school age children: Standardization, statistical analysis, psychometric properties. In *Proceedings of the 8th conference on Speech Therapy* (pp. 95-106). Athens: Ellinika Grammata (in Greek).

Vogindroukas, I., Protopappas, A., & Stavrakaki, S. (2009). *Action picture test. Test of information and grammar adequacy*. Chania: Glafki.

Voulgaraki, M. (2017). *Theory of Mind and factivity in specific language impairment: A training study.* Unpublished MA thesis Aristotle University of Thessaloniki.

Wellman, H. M., & Peterson, C. C. (2013). Deafness, thought bubbles, and theory-of-mind development. *Developmental Psychology*, *49*(12), 1-20.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception. *Cognition*, *53*, 45-57.

To cite this article: Voulgaraki, M., & Stavrakaki, S. (2017). Theory of mind and comprehension of factivity in specific language impairment: Perspectives from an intervention study. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 281-286). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Discrimination of German tense and lax vowels and their orthographic markers in monolingual and bilingual children

Amrei Walkenhorst
amrei.walkenhorst@uni-bielefeld.de

Bielefeld University

Abstract. In German there are 15 vowel phonemes that can appear in accented syllables; they appear in tenseness pairs. In the writing system, vowel phonemes are marked by a combination of a vowel grapheme and the following consonant grapheme(s). In the discussion about teaching spelling skills, it is claimed that children with German as L2 have more problems acquiring the orthographic markers for vowel tenseness. It is assumed – but has not yet been experimentally tested – that this is caused by lower discrimination skills due to interferences from the L1. To investigate the relation of discrimination ability of vowel tenseness and the acquisition of their orthographic representation, an empirically twofold study was conducted: An AX perception study was employed to test the discrimination skills of all tenseness contrasts except /ɛː ɛ/. The participants were monolingual German (G, n = 22), bilingual German-Turkish (GT, n = 37) and German-Russian (GR, n = 16) children aged 6-7. It was revealed that the bilingual children did not discriminate the items significantly less accurately than the monolingual children. However, the groups showed a significant but small difference for the /iː ɪ/ contrast: GT bilinguals discriminated the pair less accurately than the German monolinguals (G vs. GT: $z = -2.14$, $p = .03$), while the GR bilinguals did not differ significantly from the German monolingual group (G vs. GR: $z = -1.50$, $p = .13$). The ability to mark vowel tenseness in writing was further investigated by means of a qualitative longitudinal study with a subgroup of the earlier participants (n = 12). The results show that not perceptive discrimination skills but rather a (written language induced) phonological representation of the vowel system seems to be crucial for the acquisition of orthographic vowel markers. The language background does not seem to influence this.

Keywords: cross-linguistic perception, vowel tenseness, German, Turkish, Russian, acquisition of orthography

Introduction

In the vowel system of German, there are 15 vowel phonemes that can appear in accented syllables. Phonologically they appear in tenseness pairs (Hall, 2011, see below). The writing system provides only 9 vowel graphemes: <a, e, i, ie, o, u, ö, ü>. Thus, vowel tenseness is marked by the vowel grapheme in combination with the following consonant grapheme(s) (Fuhrhop, 2009): A single consonant grapheme marks the preceding vowel as a tense vowel (e.g., *hüte* [ˈhyːtə] 'hats', *beten* [ˈbeːtn̩] 'to pray'; a double consonant grapheme or two different consonant graphemes mark the preceding vowel as lax (e.g., *hütte* [ˈhʏtə] 'hut', *hüfte* [ˈhʏftə] 'hip'). Whereas the double consonant is a regular orthographic marker, lengthening <h> is irregular (e.g., *holen* [ˈhoːln̩] 'to fetch', *hohlen* [ˈhoːln̩] 'hollow'.

In the discussion about approaches to teaching spelling skills, it is claimed that children with German as L2 have more problems acquiring the orthographic markers for vowel tenseness because of lower perceptive discrimination skills due to interferences from the L1 (Becker, 2011; Bredel, 2013). So far there are only few studies that are concerned with the discrimination skills of vowel phonemes in the L2 German, and in these studies there have either not been tested all of the tenseness pairs (Darcy & Krüger, 2012) or the subjects were late grown up L2 learners (Kerschhofer-Puhalo, 2010; Nimz, 2011, 2016).

Thus in the first empirical part of the study the discrimination skills concerning vowel tenseness in all tenseness pairs except /ɛː ɛ/ were tested in German monolingual, German-Turkish and German-Russian bilingual children. There are two reasons for the choice of these contrast languages: Firstly, Russian and Turkish are phonologically different from German (see below). Secondly, both languages are spoken by a considerably big group of school children in the region of Germany where the study was conducted.

The results do not explain the observations made by Becker (2011) and Bredel (2013). Therefore, in the second empirical part a qualitative longitudinal study was conducted with a subgroup of the perception study to investigate the acquisition of the orthographic tenseness markers. The results reveal that it is not perceptive discrimination skills but rather a phonological representation of vowel tenseness that seems to be crucial for the acquisition of orthographic vowel markers. This representation of the phonological category of tenseness with its phonetic correlates of spectral features combined with duration seems to be supported by a focus on the written form. The language background does not show any influence in the participating group.

Discrimination skills of vowel tenseness in German monolingual and German-Turkish and German Russian bilingual children

The vowel systems of German, Turkish, and Russian

According to Hall (2011), German has 15 vowel phonemes that can appear in stressed syllables (/aː a/, /eː ɛ/, /ɛː ɛ/, /iː ɪ/, /oː ɔ/, /uː ʊ/, /øː œ/, /yː ʏ/, see Appendix: Figure 1), /ə/ and /ɐ/ only appear in reduced syllables as in [ˈlampə] ‘lamp’ or [ˈbɛkɐ] ‘baker’. Phonologically the vowel phonemes appear in tenseness pairs; phonetically tense vowels are longer and more peripheral than their lax counterpart. The pairs /aː a/ and /ɛː ɛ/ have the same spectral features, they differ only in duration. As in most descriptions of German phonology, *tenseness* is used in this paper to refer to the phonological category of the phonematic difference between /aː/ and /a/, /eː/ and /ɛ/, etc., not to refer to the articulatory feature of muscle strength.

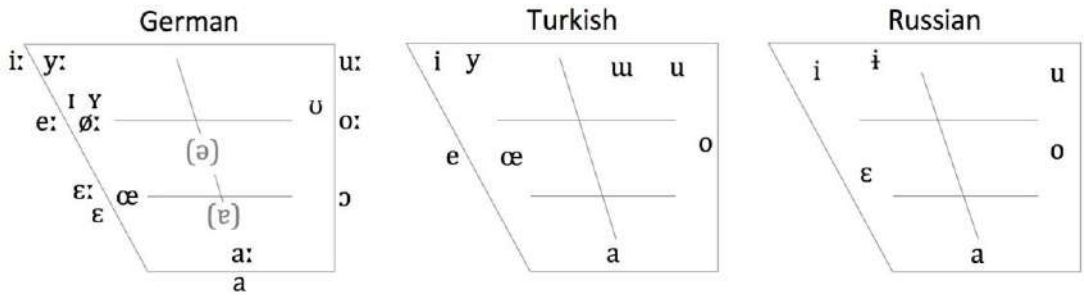


Figure 1. Vowel phonemes of German, Turkish and Russian

The vowel space of Turkish and of Russian is phonemically less differentiated (see Appendix: Figure 1). As claimed by Kornfilt (1997), Turkish has eight vowel phonemes (/i, e, a, o, u, ɯ, y, œ/); long vowels appear exclusively in loan words. Russian has six vowel phonemes that can appear in accented syllables (/ɨ, i, ɛ, a, o, u/). However, researchers discuss controversially whether [ɨ] and [i] are allophones of one phoneme or two distinctive phonemes. Vowels are subject to coarticulatory effects caused by the feature [+palatalized] of surrounding consonants that leads to fronting. In non-prominent syllables the vowel phonemes are reduced to /ʊ, ʌ, ɪ/ or /ʊ, ə, ɪ/ (Cubberley, 2002; for a discussion about the phonematic status see Iosad, 2012).

The Speech Learning Model (Flege, 1995)

As claimed by Flege (1995), the perceptive discrimination performance in an L2 depends on the perceived phonetic distance between an L2 sound and the phonetically nearest L1 sound category. If an L2 sound is perceived as phonetically different, there is a high likelihood that a new perception category is established. For example, if subjects with Turkish or Russian as L1 perceive examples of the German phonemes /uː/ and /ʊ/ as phonetically different, they will most likely establish a new category for one phoneme (and identify the second phoneme with the Turkish or Russian /u/, respectively) or even for both phonemes. In this case, the subjects show high discrimination performance for this German tenseness pair. If an L2 sound is not perceived as phonetically different, this leads to the identification of an L2 sound with the phonetically nearest L1 category. In the case of this study this would mean that bilinguals identify German tense /uː/ and lax /ʊ/ with the one corresponding category of Turkish or Russian which would lead to a low discrimination performance.

What was illustrated for the German pair /uː ʊ/ would be expected for all tenseness contrasts that correspond phonetically to only one Turkish or Russian vowel category. However, in Russian, the front rounded area of the vowel space is not phonemically (but according to Kouznetsov 2001 at least allophonically) used. It can be assumed that the German phonemes /øː, œ, yː, ʏ/ are perceived as phonetically different from the nearest Russian vowel categories. Since there are four 'new' German phonemes, on the other hand, lower discrimination performance might be expected compared to German monolinguals.

Experimental design

An AX-perception task was applied with pseudo words. All tenseness pairs were tested in the two contexts [ˈfVpə] and [ˈgVpə]. The ISI was 500 ms. To keep the task in a reasonable length for first-graders, one contrast pair was excluded; /ɛː ɛ/ was not tested since in many German speaking regions /ɛː/ matches phonetically with /eː/ (Maas, 2006). To make the experiment suitable for children, the study was used in the context of a story about two frogs: One frog wants to learn the frog language. The children were asked to judge whether he repeated his uncle's words successfully or not. The participants were shown two animated frogs that opened and closed their mouths one after another while the stimuli were presented.

Participants

The participants were 22 German monolinguals (G), 37 German-Turkish (GT) and 16 German-Russian bilinguals (GR) at the age 6-7. The study was conducted during the first half year of school enrolment. The parents were asked to fill out a questionnaire concerning the use of language(s) in the families, e.g., Which language(s) is/are how often spoken in the family? Was your child born in Germany, if not, how old was it when you migrated to Germany? The answers showed that for most bilingual children who could be acquired for the present study, German was not a clear L2, rather they were surrounded by both languages from the beginning on.

Results and discussion

The graph on the left in Figure 2 shows the mean accuracy of the language groups. For the accuracy, only answers to *different* items were taken into consideration. (The data of children who rated an unusual high number of *same* items as different were excluded. This concerned four children who rated between 8 and 28 of the 28 *same* items as different.) An accuracy of .8 means that 80% of the *different* items were identified as being different.

A generalized linear mixed model (GLMM) was fit to the binomial data, with LANGUAGE as a fixed and PARTICIPANT as a random factor. It was revealed that the bilingual groups do not differ significantly from the G monolingual group (G vs. GT: $z = .53$, $p = .60$, G vs. GR: $z = .48$, $p = .37$). However, more detailed inspection into the different vowel categories revealed that the groups show different results for the /iː ɪ/ contrast (see second graph in Figure 2): GT bilinguals discriminated the pair significantly, but only slightly less accurately than the G monolingual group (G vs. GT: $z = 2.14$, $p = .03$), while the GR bilinguals did not differ significantly (G vs. GR: $z = -1.50$, $p = .13$). This could be explained by phonetic or phonological interference from Turkish and Russian. In Turkish there is only one vowel phoneme in the front closed unrounded vowel space (/i/), whereas in Russian this same region is divided in two phonemes (/i, ɨ/, see Appendix: Figure 1). While there are no further significant differences between the language groups, the discrimination performance of all three language groups varies with regard to the vowel contrasts. The accuracy in the /aː a/ and the /uː ʊ/ contrasts is around chance level. This could be due to the stimuli used in the experiment. In the case of the /aː a/ contrast another explanation could be that the vowels only differ in duration, not in spectral features.

In relation to the Speech Learning Model the results can be read as follows: Except for the /iː ɪ/ contrast the results cannot be interpreted as confirmation of the hypothesis that German tense and lax vowels are identified with one Russian or Turkish vowel phoneme. The results imply, but do not statistically prove, that bilinguals developed perception categories for the German tense and lax vowels that lead to a discrimination performance that is quite similar to that of G monolinguals.

How can the results be explained regarding the observation that children with German as L2 have more problems acquiring orthographic tenseness markers? One explanation could be that this observation was made for children with German as L2 (see above). For the participants that could be acquired for the present study, though, German is rather a second L1 than an L2 and thus is can be assumed that their perceptive discrimination skills differ from those of successive bilinguals. However, the data of the second part of the study imply that there is another factor besides or maybe even instead of discrimination skills, as will be shown in the following section.

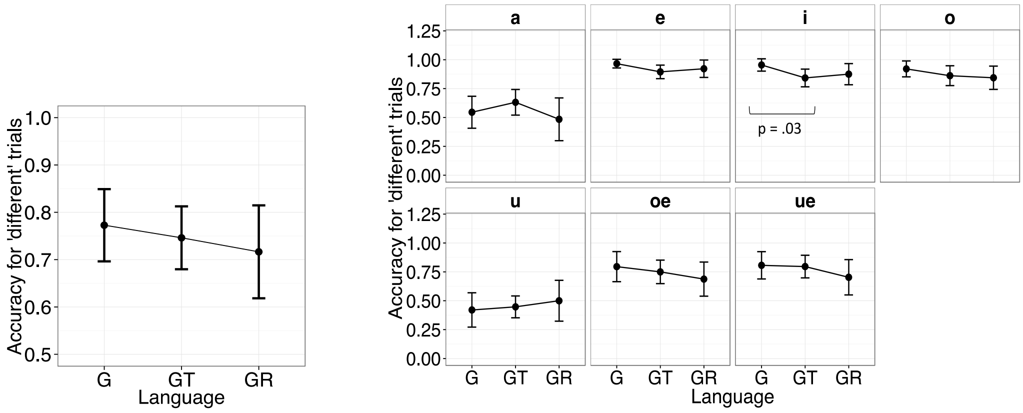


Figure 2. Mean Accuracy of German monolinguals (G), German-Turkish (GT) and German-Russian (GR) bilinguals (left), differentiated by vowel contrast (right)

The acquisition of orthographic vowel tenseness markers in monolingual and bilingual children

A qualitative longitudinal study was conducted with a subgroup of the participants of the perception experiment (n = 12). During the first two years after school enrolment, the following data were collected: the development of the perceptive discrimination performance, worksheets and texts from writing tasks in class, a standardized test of phonological awareness (Stock, Marx, & Schneider, 2003), writings of pseudo words similar to the stimuli of the perception experiment as well as audio and video recordings of interactions between the children about spelling and phonological features of words.

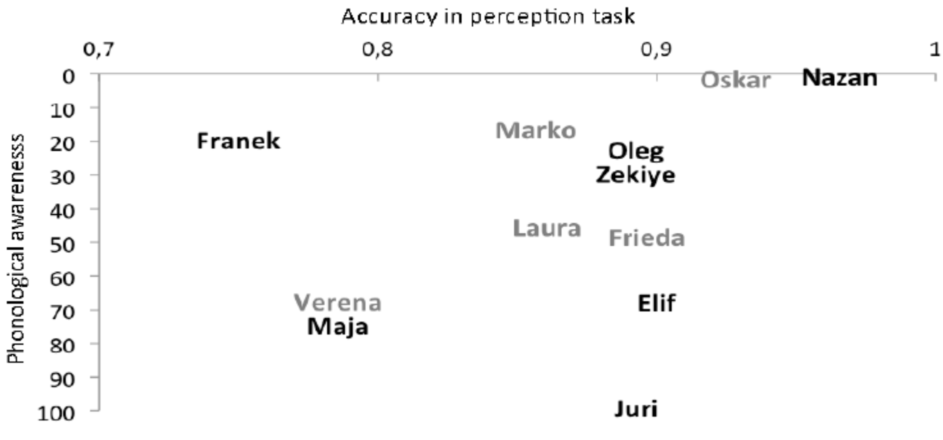


Figure 3. Accuracy in perception task and result in test of phonological awareness in percentage rate at the end of 1st grade; names of monolinguals in grey

Phonological awareness, writing skills and discrimination ability

Figure 3 shows the accuracy the children achieved in the discrimination experiment during the first half year of 1st grade and the results of the test of phonological awareness (in percentage rate) at the end of 1st grade. A comparison of the test results with the level of writing skills shows that the level of phonological awareness reflects the writing skills of the children. Since the test is working exclusively on a segmental, phonematic level, it is most probably testing the same skills as writing, just in another medium. The table shows no pattern in relation to the linguistic background of the participants. Two subgroups show opposite results in the discrimination task and the test of phonological awareness: Nazan and Oskar perform very well in the perception task, whereas their phonological awareness is very low. Verena, Maja and Juri, on the other hand, reach high percentage rates on the test of phonological awareness and Verena and Maja show relatively low discrimination accuracy in comparison to the other participants of the longitudinal study. Interestingly, the children of these two subgroups do not only show similarities in these two domains, but also in their writing strategies.

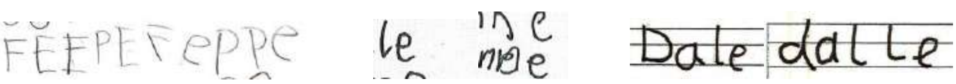
Writing strategies

Until the end of 2nd grade, Nazan and Oskar do not mark different vowel tenseness at all or only by their choice of the vowel grapheme. For example, Oskar writes *Pite* for [ˈpiːtə] and *Peth* for [ˈpɪtə] (see Figure 4a). In the recordings the children repeat the words slowly. By lengthening [ɪ], [ʏ], and [ʊ] they articulate [eː], [øː] and [oː], respectively, and write the corresponding grapheme (see Figure 5b). Nazan also shows a confusion of vowel phoneme and grapheme in other vowels (<o> for [uː] and [ʊ], <e> for [øː], see Figure 4c). The recordings indicate as well that the children analyse words phonetically (not phonologically, as would be required by the writing system): Oskar claims to hear [h] in several words at positions where this sound is not distinctive in German, e.g., as aspiration after a voiceless stop (graphically marked in *Peth*) or at the word-final position in [ˈgaːpə] (*Gapeh*, see Figure 4c).

a) [ˈpiːtə] [ˈpɪtə] b) [ˈpyːtə] [ˈpʏtə] c) [ˈguːpə] [ˈgʊpə] d) [ˈgaːpə]

Figure 4. Pseudowords written by Oskar (a, d) and Nazan (b, c)

Verena, Maja, and Juri, on the other hand, mark vowel phonemes at the end of 2nd grade mostly by choice of the vowel grapheme in combination with the following consonant grapheme(s) (see Figure 5a and b). During that time they produce almost exclusively writings that mark phonological (not phonetic) segments. They use orthographic markers according to the writing system much earlier in comparison to the other participating children, even in loanwords (e.g., *clown* and *pool*). Since loanwords are usually imported together with the spelling, this means that the children memorized the written form. In the recordings, Verena, Maja and Juri show an orientation towards the written form, as well. They recode and compare the written and the phonetic or phonological forms of the words they wrote. In some situations they refer to a written form they memorized and transfer it to a pseudo word that they want to write down. During the writing process they analyse the phonetic form without manipulating it.



a) [ˈfeːpə] [ˈfɛpə] b) [ˈkøːtə] [ˈkœtə] c) [ˈdaːlə] [ˈdalə]

Figure 5. Pseudowords written by Verena (a) and Maja (b) and Juri (c)

As mentioned above, the phonetic correlates of the phonological category of tenseness are vowel duration and qualitative features (see above). From a phonetic point of view, some German tenseness pairs are ambiguous. On the one hand, /iː/ and /ɪ/, /yː/ and /ʏ/, /uː/ and /ʊ/ pair up *phonologically*, which is also reflected by the writing system: Both vowels of a tenseness pair are represented by the same grapheme (<i, ü, u>, respectively; actually <ie> is the regular grapheme for /iː/, however, due to the pedagogical assumption that this might be too difficult for children, they are often taught <i> as a regular grapheme for /iː/). On the other hand, /ɪ/, /ʏ/, and /ʊ/ are *phonetically* much closer to /eː/, /øː/ and /oː/, because they share similar qualitative features (Sendlmeier, 1981; Sendlmeier & Seedbode, 2006).

Since Nazan and Oskar tend to neutralize vowel duration, they focus on the qualitative vowel features only and thus do not analyse the distinctive feature of vowel tenseness as the combination of qualitative features *and* duration. Verena, Maja and Juri do not show this reduction of vowel contrasts to spectral features. They seem to develop a phonological vowel system with adequate tenseness pairs already shortly after school enrolment. It can be assumed that this is due to their orientation towards written forms that mark the phonetically ambiguous tenseness pairs.

Conclusion

By means of a phonetic perception study the discrimination ability of German tense and lax vowels was tested in German monolingual and bilingual school children. The results reveal that the bilingual groups do not differ from the monolingual group significantly. Only one significant but very small difference was found for one vowel contrast. Besides, the results reveal that also German monolingual children do not differentiate the /aː a/ and /uː ʊ/ contrast better than chance level. Thus, the results do not explain the observed difficulties of children with German as L2 in the acquisition of orthographic tenseness markers. The results of the qualitative longitudinal study show that high perceptive discrimination skills do not automatically lead to a successful acquisition of the graphic markers of vowel tenseness: Nazan and Oskar show high perceptive discrimination skills and low orthographic proficiency, including orthographic tenseness markers, whereas Juri, Maja and Verena show contrary results. Instead of perceptive discrimination skills, the children’s writing strategy seems to be crucial for the acquisition of the vowel tenseness markers: While Nazan and Oskar focus on phonetic features which they mark graphically, repeating words slowly with a neutralised vowel duration, the children of the second subgroup are focusing on written forms, memorising, analysing, and comparing them with each other.

An explanation for these contrary developments of the two subgroups could be differing representations of the vowel system the children rely on while writing and reading: As mentioned above, the phonetic correlates of the phonological category of tenseness are duration and spectral features. Some tenseness pairs are phonetically ambiguous. In the writing system, however, the tenseness pairs are marked, since a tense and a lax vowel phoneme is represented by the same vowel grapheme. It seems that Nazan’s and Oskar’s high discrimination skills combined with their focus on phonetic forms impede the development of a representation of the vowel system with the phonological category of tenseness as a combination of qualitative *and* quantitative phonetic features. Instead, their representation of vowels consists of categories that are based on spectral features only. The writing strategies of Juri, Maja and Verena, on the other hand, imply that they develop a phonological representation of the vowel system with adequate tenseness pairs, however, not due to their perceptive discrimination skills, but rather by focussing on written forms where the phonetically ambiguous tenseness pairs are graphically marked. The findings have relevant implications for the teaching methods of orthography. In many schools in Germany, including the classes of the participating children, a teaching method is used where the analysis of the spoken word represents the basic activity children are supposed to use to derive the spelling. These methods are based on the wrong assumption that the phonetic form represents the written form directly. Apparently some children cannot benefit from that, as the profiles of Nazan and Oskar clearly show. This is no surprise since it is questionable how anything can be learned about the writing system by analysing the phonetic form (Birk & Häffner, 2005). Instead, methods should be used that are based on a concept of written language as a system that represents phonological, syllabic and morphological information and where thus written forms are focused on.

References

Becker, T. (2011). *Schriftspracherwerb in der Zweitsprache. Eine qualitative Längsschnittstudie.* [Acquisition of writing in the second language. A qualitative longitudinal study]. Baltmannsweiler: Schneider-Verlag Hohengehren.

Birk, E., & Häffner, S. (2005). Was ist phonologische Bewusstheit? Schrifttheoretische Analyse einer psychologischen Fragestellung. [What is phonological awareness? A theoretical graphematical analysis of a psychological question]. In H.-W. Huneke (ed.), *Geschriebene Sprache. Strukturen, Erwerb, didaktische Modelle* (Vol. 45, pp. 53-72). Heidelberg: Mattes.

Bredel, U. (2013). Silben und Füße im Deutschen und Türkischen - Orthographieerwerb des Deutschen durch türkischsprachige Lerner/innen. [Syllables and foot in German and in Turkish – Acquisition of German orthography in Turkish-speaking learners]. In A. Deppermann (ed.), *Das Deutsch der Migranten* (pp. 369-390). Berlin: De Gruyter.

Cubberley, P. V. (2002). *Russian: A Linguistic Introduction*. Cambridge/New York: Cambridge University Press.

Darcy, I., & Krüger, F. (2012). Vowel perception and production in Turkish children acquiring L2 German. *Journal of Phonetics*, *40*(4), 568-581. https://doi.org/10.1016/j.wocn.2012.05.001

Flege, J. E. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (ed.), *Speech perception and linguistic experience. Issues in cross-language research* (pp. 233-277). Baltimore, MA: York Press.

Fuhrhop, N. (2009). *Orthografie* [Orthography] (3rd ed.) Heidelberg: Winter.

Hall, T. A. (2011). *Phonologie: Eine Einführung* [Phonology. An Introduction.] (2nd ed.) Berlin: De Gruyter.

Iosad, P. (2012). Vowel reduction in Russian: No phonetics in phonology. *Journal of Linguistics*, *48*(3), 521-571.

Kerschhofer-Puhalo, N. (2010). Preferences in L2-Vowel Perception. *New Sounds. 10th International Symposium on Acquisition of Second Language Speech* (pp. 245-250).

Kornfilt, J. (1997). *Turkish*. London/New York: Routledge.

Kouznetsov, V. (2001). Spectral Dynamics and Classification of Russian Vowels. *XI Session of the Russian Acoustic Society* (pp. 439-442).

Maas, U. (2006). *Phonologie: Einführung in die funktionale Phonetik des Deutschen* [Phonology: An Introduction to Functional Phonetics] (2nd ed.) Göttingen: Vandenhoeck & Ruprecht.

Nimz, K. (2011). Vowel Perception and Production of Late Turkish Learners of L2 German. *International Congress of Phonetic Science*, *17*, 1494-1497.

Nimz, K. (2016). *Sound perception and production in a foreign language. Does orthography matter?* Potsdam: Universitätsverlag Potsdam. Retrieved from https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/8879 (27.10.2017)

Sendlmeier, W. F. (1981). Der einfluß von qualität und quantität auf die perzeption betonter vokale des deutschen. [The influence of quality and quantity on the perception of accented vowels in German], *Phonetica*, *38*, 291-308.

Sendlmeier, W. F., & Seebode, J. (2006). Formantenkarten des deutschen Vokalsystems. [Formant maps of the vowel system of German.] Retrieved from https://www.kw.tu-berlin.de/fileadmin/a01311100/Formantkarten_des_deutschen_Vokalsystems_01.pdf (27.10.2017)

Stock, C., Marx, P., & Schneider, W. (2003). *BAKO 1-4. Basiskompetenzen für Lese-Rechtschreibleistungen. Ein Test zur Erfassung der phonologischen Bewusstheit vom ersten bis vierten Grundschuljahr*. [Basic skills for proficiency in reading and writing. A test measuring phonological awareness from 1st to 4th grade of elementary school.] Göttingen: Beltz Test.

To cite this article: Walkenhorst, A. C. (2017). Discrimination of tense and lax vowels and their orthographic markers in monolingual and bilingual children. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 287-293). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Schwa realization in French: Investigations using automatic speech processing and large corpora

Yaru Wu[1], Martine Adda-Decker[1,2]
yaru.wu@univ-paris3.fr, madda@limsi.fr

[1]UMR7018, CNRS-Sorbonne Nouvelle, [2]LIMSI, CNRS, Université Paris-Saclay

Abstract. This study aims to analyze differences in schwa production according to phonological factors, speaking styles, sex and socio-professional categories. Three large corpora of public journalistic speech (ESTER, 1M word-tokens; ETAPE, 0.5M word-tokens) and casual speech (NCCFr, 0.4M word-tokens) are explored with a method based on automatic speech processing. Schwa realization is automatically decided via forced alignment, which has a successful performance rate of 95%. In order to determine words with potential schwa, we used the phonological transcription of Lexique380 as a reference. Only words of type CəCV (e.g., *semaine*/səmɛn/) and type CəCCV (e.g., *retraite*/rətrɛt/), i.e. including a potential schwa in word-initial position, are chosen to better control word structure and position related variability. Linear mixed model are used for the statistical analyses. Know effects could be replicated with the proposed method. A clear effect of speech style is found on schwa realization (present/absent). Gender differences are also found: male speakers tend to have higher schwa deletion rates than females. Interestingly, journalists tend to delete more schwa than politicians according to our data. Methods and tools developed for automatic speech recognition may be very helpful for investigating linguistic hypotheses in production. As they may readily apply to additional data sets, they offer the opportunity to improve and fine-tune the current descriptions on schwa deletion in French.

Keywords: schwa, large corpora, speech style, forced alignment

Introduction

As one of the most complicated phenomena in French Phonology, schwa (an unstable vowel that alternates with ∅) realization has been studied for more than a century. According to Grammont (1894), the presence and absence of schwa depends on its consonantal contexts: schwa becomes mandatory when the surface form resulting from schwa deletion has three or more consonants in a row (Grammont's Three Consonants Rule). Since then, numerous phoneticians and phonologists have studied the phenomena (see for example in Côté, 2000) and Durand, Laks, and Lyche (2009)). We intend to investigate the phenomenon using tools from automatic speech processing, which allow us to study phonological phenomena using huge data collection. In this study, both phonological and socio-linguistic factors are analyzed. Prosody influence is not analyzed since that there is not yet enough prosodic annotation in our data. We propose a new approach for linguistic analyses using tools from automatic speech processing.

Corpora

We make use of three publicly available large corpora, which correspond to three different production situations. Two journalistic speech corpora and a casual speech corpus are used for our analyses: ESTER (Galliano, Geoffrois, Gravier, Bonastre, Mostefa & Choukri, 2006), ETAPE (Gravier, Adda, Paulson, Carré, Giraudel, & Galibert, 2012) and NCCFr (Torreira, Adda-Decker, & Ernestus, 2010). ESTER corpus contains 100 hours of radio broadcast news shows; the ETAPE corpus consists of 13.5 hours of radio data and 29 hours of TV data, containing mainly debates and free conversations; NCCFr is composed of 35 hours of casual French conversation between friends.

Method

Speech alignment

Speech data are automatically segmented and labeled using the LIMSI speech transcription system (Gauvain, Lamel, & Adda, 2002) in forced alignment mode. Forced alignment consists in finding the optimal association between speech and acoustic HMM phone models according to the phonemic transcription as specified by the speech's word level transcription and a pronunciation lexicon. In forced alignment mode, the transcription system is constrained (or forced) to use the proposed manual transcription to find the optional association between the speech signal and the corresponding sequence of HMM phone models, whereas in normal transcription mode, any phone sequence corresponding to any admissible word sequence may be explored (Adda-Decker & Lamel, 2000). In order to move from a phonemic transcription to a more phonetic one (i.e. which may adapt to the speakers' production), pronunciation variants can be added to the pronunciation lexicon of the transcription system. As a result, the forced alignment system remains constrained for the word sequence but is free to choose among the most suitable variants in order to stay as close as possible to the speaker's pronunciation. As an output, the system provides word and phone boundaries, along with labels and boundaries for pause, hesitation and breathe. For our study focusing on the realization and non-realization of schwa, schwa variants were introduced into the transcription system: each word containing a schwa in its pronunciation is recursively added a variant without schwa. Thus, schwa realization can be automatically decided via forced alignment (Figure 1). A subset of the segments concerning schwa was verified manually. The successful performance rate of the automatic alignment concerning schwa realization is 95%.

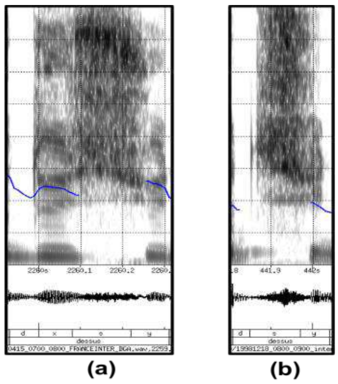


Figure 1. Examples of variant selection and phone segmentation of the word *dessus* (/dəsy/, above) produced by the same speaker

Data preparation

We chose to analyze polysyllabic words with schwa in word initial position, i.e. CəCV (e.g., *semaine* /səmɛn/ 'week') and CəCCV (e.g., *retraité*/rətrɛte/ 'retired') words to control for phoneme position related variability. Note that CCəCV words (e.g., *prenons*/prənɔ̃/ '2nd person plural formof 'take') were discarded, as there were only few occurrences and most importantly very few in types. Corresponding information (i.e. aligned pronunciation, contexts of the word in question, corpus name, recording sessions, speaker, sex and socio-professional status) is extracted for each word in question. Speakers with fewer than 50 occurrences of word-tokens are discarded in this study, leaving 158 speakers (70 speakers in the ESTER corpus, 42 in the ETAPE corpus and 46 in the NCCFr corpus) to be included in this study.

In order to obtain information on the absence and presence of schwa, we used Lexique380 (New, Brysbaert, Veronis, & Pallier, 2007) as a reference. Thus, we are able to quantify the absence and presence of schwa by comparing the full pronunciation forms (Lexique380) to the aligned pronunciation forms (production of speakers aligned by the LIMSI transcription system) and categorize word tokens into 'schwa present' and 'schwa absent'. Information concerning the nature of the consonants (voiceless stop 'o', voiced stop 'O', voiceless fricative 'f', voiced fricative 'F', nasal

'N', liquid 'L', glide 'G') in each word is generated for our statistical analyses. Vowels are annotated as 'V' and schwas are annotated as 'ə' (unchanged).

Important factors

Our corpora allow us to study both phonological and socio-linguistic factors. With regard to phonological factors, we are particularly interested in the number of consonants in the surface form resulting from schwa deletion ('2C' for CəCV and '3C' for CəCCV) and the post-lexical context (i.e. cross-word left contexts: pause 'O#_' vs. consonant 'C#_' vs. vowel 'V#_') of the word in question. We suppose that schwa tends to be deleted more in word form '2C' (CəCV) than in word form '3C' (CəCCV) and in words with context 'V#_' than in those with context 'C#_'. Côté (2000) and many others observe that schwa after a post-pausal position (e.g., *demande-la*/dəmãd la/ 'request it') is optional. We are interested in finding out how the context pause 'O#_' would influence schwa realization in our data.

Speech styles, sex and socio-professional status are chosen as our socio-linguistic factors. We suppose that schwa deletion occurs more frequently in NCCFr (casual 'private' speech) than in ETAPE (casual journalistic public speech) and ESTER (formal journalistic speech). As for ESTER and ETAPE, we expect fewer schwas in ETAPE than in ESTER. Simpson (2000) and many others show that male speakers articulate with less precision than female speakers. We intend to analyze whether male speakers delete schwa more than female speakers. Analyses on the influence of social-professional status are limited to the ETAPE corpus to better control for speech style related variation. We are particularly interested in differences between journalists and politicians on schwa realization. We are interested in finding out whether journalists produce fewer schwas than politicians, since journalists tend to be more aware of limited time slots for their speech.

Results

Token analyses on phonological and socio-linguistic factors

Factors listed in the previous section are analyzed by crossing different phonological factors and by crossing different socio-linguistic factors. The influence of socio-professional status is shown in a separate figure since it concerns the ETAPE corpus only.

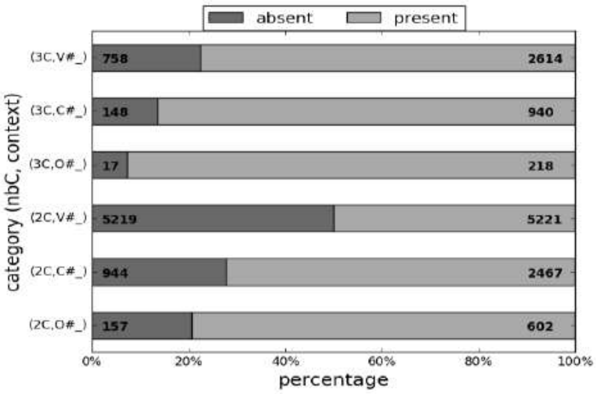


Figure 2. Horizontal bar plot providing a combined view of the percentages of deleted and realized schwas for different conditions (Y-axis): number of consonants in the surface form (2C vs. 3C) and post-lexical contexts (V#_ vs. C#_ vs. O#_). The number of occurrences per category is also displayed

Figure 2 illustrates the influence of phonological factors on schwa realization in several contrasting conditions: the number of consonants in the surface form and post-lexical contexts. From top to bottom, three '3C' conditions specifying the three left post-lexical contexts (V#_, C#_ and O#_) are shown, followed by the similar three '2C' conditions. As predicted by Grammont's Three Consonant Rule, schwa tends to be more present when it concerns '3C'. As expected, schwa deletion rates are the highest in left vowel context (V#_). The context pause ("O#_") and the context consonant seem to have a similar influence on schwa realization.

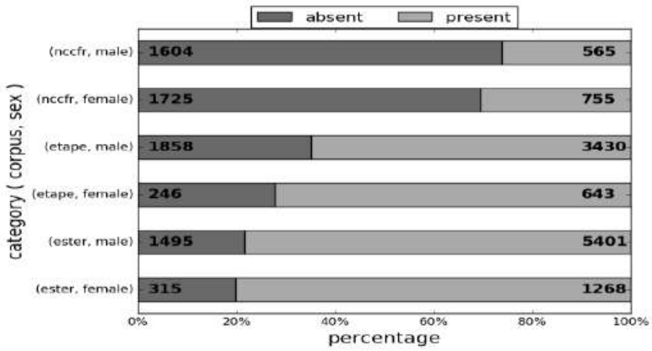


Figure 3. Horizontal bar plot providing a combined view of the percentages of deleted and realized schwas for different conditions (Y-axis): speech style/ corpus (ESTER vs. ETAPE vs. NCCFr) and sex (male vs. female)

Figure 3 shows the influence of speech style and sex on schwa realization. Schwa tends to be deleted more in NCCFr than in ESTER and ETAPE; schwas are deleted more in ETAPE than in ESTER. Females are shown to produce slightly more schwas than males.

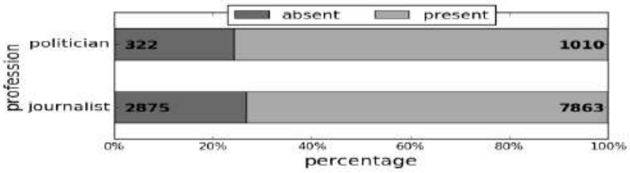


Figure 4. Horizontal bar plot providing a combined view of the percentages of deleted and realized schwas for the profession of the speakers (journalist vs. politician)

The influence of profession is shown in Figure 4. Journalists tend to produce fewer schwas than politicians. However, the difference between the two categories is rather small.

Model-based analyses

To validate token analyses in previous section while taking into consideration the inter-speaker and other data-dependent variability, generalized linear mixed models are used for the analyses of this study (McCulloch & Neuhaus, 2001) using the package lme4 (Bates, Mächler, Bolker, & Walker (2014) in R (Team 2014). The effects of number of consonant (in the surface form resulting from schwa deletion), post-lexical context, speech style and sex are considered within the same model. To avoid collinearity related problem, we conducted a separate model for the effect of profession.

We used 'number of consonants in surface form ('2C' for CəCV and '3C' for CəCCV, reference: CəCCV)', 'post-lexical context (O#_, C#_ or V#_, reference: O#_)', 'corpus/speech style (i.e. ESTER, ESTER or NCCFr, reference: ESTER)', and 'sex (male or female, reference: female)' as fixed effect. The following random terms were considered in the model: a random intercept per speaker and one per consonant nature; by-speaker slopes for the effect of type and context; by-consonant-nature slopes for the effect of context and gender. Model-based post-hoc tests were performed for fixed effects to get information on each level of each fixed effect that has more than two levels.

Results show that the realization of schwa is under the influence of consonants in the surface form resulting from schwa deletion, post-lexical context, speech style and sex. Fewer schwas are observed when it concerns 2C than when it concerns 3C (log odds ratio = -0.719, $|Z|=3.528$, $p<0.001$). Both C#_ and V#_ contexts have a significantly negative effect on the realization of schwa with respect to that observed in O#_ (C#_: log odds ratio = -0.834, $|Z|=2.560$, $p<0.05$; V#_: log odds ratio = -1.887, $|Z|=5.905$, $p<0.001$). The probability to observe a schwa decreases significantly both in ETAPE (log odds ratio = -0.546, $|Z|=4.416$, $p<0.001$) and in NCCFr (log odds ratio = -2.798, $|Z|=20.499$, $p<0.001$)

with respect to that observed in ESTER. Finally, female speakers produce fewer schwas than male speakers (log odds ratio = -0.339, |Z|=2.445, $p<0.05$).

Figures 5a and 5b show complementary information on different levels of fixed effects that have more than two levels (i.e. the effect of post-lexical context and that of speech style/corpus). The overlap of the error bars on the figures doesn't necessarily indicate the non-significance between the variables, as the standard error of a model coefficient differs from that of a model's predicted value (Fox, 2003).

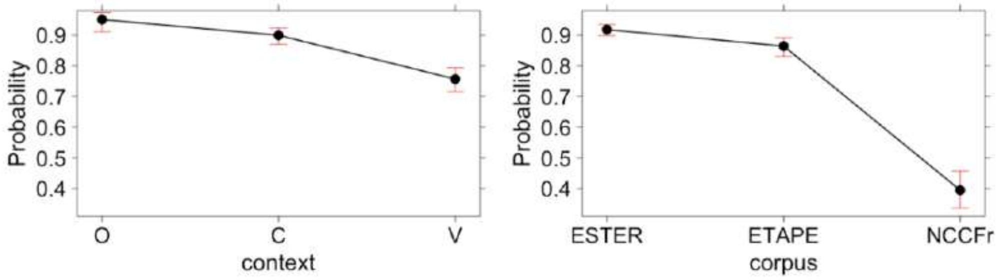


Figure 5. Probability of the presence of schwa for the effect of the post-lexical context (a) and the effect of corpus/ speech style (b). (a) Post-lexical context: word final vowel (V#_) or final consonant (C#_) or a pause (O#_); (b) Corpus/speech style context: ESTER (formal journalistic speech) vs. ETAPE (casual journalistic speech) vs. NCCFr (causal conversational speech)

As shown in Figure 5(a), post-lexical context influences schwa deletion. Schwa tends to be deleted more when the preceding word ends in vowel (V#_); schwas in words preceded by a consonant are prone to be preserved since the more consonants in the surface form, the less schwa tends to be omitted (Grammont's Three Consonants Rule). Interestingly, the context O#_ has the lowest schwa deletion rate. Post-hoc test based on the model shows that the influence of the context V#_ is significantly different from that of the context C#_ ($p<0.001$); the context V#_ is significantly different from that of the context O#_ ($p<0.001$). As mentioned earlier, schwa is considered optional in the literature when it concerns the context O#_. Our results show that word-initial schwas in polysyllabic words tend to be preserved when the post-lexical context is O_#. Figure 5(b) suggests that schwa tends to be preserved more in corpus ESTER than in corpus ETAPE ($p<0.001$) and it tends to be deleted more in NCCFr than in ETAPE ($p<0.001$). Our results concerning speech style show that the less formal the speech style, the less probable schwa is realized in the speakers' production.
Furthermore, we have investigated the effect of profession using ETAPE corpus only and we are particularly interested in 'journalists' and 'politicians'. We considered profession (journalist vs. politician, reference: journalist) as fixed effect. The following random terms were included in the model: a random intercept per speaker and one per consonant nature; by-speaker slopes for the effect of profession; by-consonant-nature slopes for the effect of profession. The probability to observe a schwa increases significantly for politicians (log odds ratio = 0.625, |Z|=2.284, $p<0.05$) with respect to that observed for journalists. Results concerning the effect of profession indicate that politicians tend to produce schwa more often than journalists.

Discussion and conclusions

Our results on the number of consonants in the surface form resulting from schwa deletion ('2C' for CəCV and '3C' for CəCCV) and the extended research on post-lexical context (here, context C#_) are consistent with Grammont's Three Consonants Rule. We tend to observe fewer schwas for words with 2C than for words with 3C. Words preceded by a vowel (V#_) tend to favor schwa deletion; words preceded by a consonant (C#_, 3 or 4 consonants in a row in the surface form if schwa is deleted) or a pause are less likely to have their schwa deleted. Our results concerning phonological factors (number of consonants, post-lexical context) show that schwa realization depends not only on the word itself, but also on its preceding context. Socio-linguistic factors influence schwa realization as well: schwa is deleted more often in less formal settings. Schwa is more likely to appear in journalistic speech than in conversational speech. Male speakers tend to produce fewer schwas than female speakers. Public speakers have different speaking strategies with respect to schwa realization: journalists tend to produce fewer schwas than politicians. Our results contribute to fine-tune the current descriptions on schwa deletion in French. The proposed method for phonetic and phonological analyses using speech technology, in particular forced alignment with pronunciation variants, helps explore and test various linguistic hypothesis efficiently in large speech corpora.

Appendix

Examples of pronunciation references and aligned pronunciations:

	Pronunciation reference (Lexique380)	Aligned pronunciation
semaine (week)	səmɛn	səmɛn
		smɛn
retraité (retired)	rətrɛte	rətrɛte
		rtrɛte

Acknowledgements

The present study was partially funded by the French Investissements d'Avenir - Labex EFL program (ANR-10-LABX-0083).

References

Adda-Decker, M., & Lamel, L. (2000). The use of lexica in automatic speech recognition. In *Lexicon Development for Speech and Language Processing* (pp. 235-266). Springer Netherlands.

Adda-Decker, M., Nemoto, R., & Durand, J. (2009). Stratégies de démarcation du mot en français: une étude expérimentale sur grand corpus. *Actes des Proceedings of JEL*, 91-96.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Côté, M. H. (2000). *Consonant cluster phonotactics: A perceptual approach* (Doctoral dissertation).

Durand, J., Laks, B., & Lyche, C. (2009). Le projet PFC (phonologie du français contemporain): une source de données primaires structurées. *Phonologie, variation et accents du français*, 19-61.

Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of statistical software*, *8*(15), 1-27.

Galliano, S., Geoffrois, E., Gravier, G., Bonastre, J. F., Mostefa, D., & Choukri, K. (2006). Corpus description of the ESTER evaluation campaign for the rich transcription of French broadcast news. In *Proceedings of LREC* (Vol. 6, pp. 315-320).

Gauvain, J. L., Lamel, L., & Adda, G. (2002). The LIMSI broadcast news transcription system. *Speech communication*, *37*(1), 89-108.

Grammont, M. (1894). Le patois de la Franche-Montagne et en particulier de Damprichard (Franche-Comté). IV: La loi des trois consonnes. *Mémoires de la Société de linguistique de Paris*, *8*, 53-90.

Gravier, G., Adda, G., Paulson, N., Carré, M., Giraudel, A., & Galibert, O. (2012). The ETAPE corpus for the evaluation of speech-based TV content processing in the French language. In *LREC-Eighth international conference on Language Resources and Evaluation*.

McCulloch, C. E., & Neuhaus, J. M. (2001). *Generalized linear mixed models*. John Wiley & Sons, Ltd.

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied psycholinguistics*, *28*(4), 661-677.

Simpson, A. P. (2000). Gender-specific differences in the articulatory and acoustic realization of interword vowel sequences in American English. In *5th Seminar on Speech Production: Models and Data. Kloster Seeon* (pp. 209-212).

Swerts, M., & Krahmer, E. (2010). Visual prosody of newsreaders: Effects of information structure, emotional content and intended audience on facial expressions. *Journal of Phonetics*, *38*(2), 197-206.

Team, R. C. (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.

Torreira, F., Adda-Decker, M., & Ernestus, M. (2010). The Nijmegen corpus of casual French. *Speech Communication*, *52*(3), 201-212.

To cite this article: Wu, Y., & Adda-Decker, M. (2017). Schwa realizations in French: Investigations using automatic speech processing and large corpora. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 294-300). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Lusophones' pronunciation of Czech and its perception by native Czech speakers

Eliška Zaepernicková[1], Martin Havlík[2], Barbora Kraftová[3]
eliska.zaepernickova@gmail.com, havlik@ujc.cas.cz, bkraftova@gmail.com

[1]Charles University, [2]Czech Academy of Sciences, [3]Instituto Camões em Praga

Abstract. The role of a speaker's L1 seems to be neglected when the perception of non-native accented speech is addressed. In our study, we aim to show how a presupposed L1 may bias the perception of a speaker's accentedness. In a perceptual test, listeners with L1 Czech were given 52 speech samples from a total of 26 speakers. Out of these 10 were native speakers of Portuguese, both the Brazilian and European variety. The listeners were asked to fulfil 2 tasks: (1) assign the speaker to a predefined language group and (2) evaluate the speaker's accentedness. The results of the perceptual test show that Portuguese native speakers who were rated as being at least intermediate (5 points on a 10-point scale) are constantly perceived as being native speakers of Russian, Ukrainian or another Slavic language. The results also suggest there is a difference between female and male perception, females are perceived as being more Russian or Ukrainian than men. Furthermore, we accompany the results of the perceptual test with a corpus-based study of connotations of different foreign accents in the Czech media. These suggest that Russian, Ukrainian and Slovak accented speech are mostly addressed when speaking of perpetrators of crime. When compared with the results of our perceptual test, these results of the corpus-based study suggest that Portuguese native speakers may be biased, just based on their presupposed L1 which is consistently perceived as being of Slavic origin by Czech listeners.

Keywords: non-native accent, perception, Czech, Portuguese, covert prestige

Introduction

Second language acquisition (SLA) research in today's globalized world should go far beyond its primary focus on accentedness, intelligibility and comprehensibility. While we do not want to undermine the role of L2 production research, we aim at emphasizing the importance of research which focuses mainly on the perception of non-native accented speech and how this perception is connected to the L1 of both, speakers and listeners. To our knowledge, not a lot has been done so far in the research on the perception of non-native accented speech. Such research is important for two reasons: firstly, it helps to deepen our knowledge about the existing phonological systems of different languages, and secondly, the evaluation of non-native speakers' accentedness by native listeners, as well as their attitude towards their speech is based on the presumed L1 of the speakers. The latter has been the topic of studies focusing on L2 English, such as Lev-Ari and Keysar (2010) who focused on the credibility of non-native accented speech and showed that accented speech is perceived as less true than the non-accented one. Similarly, Rubin (1992) extended the research of non-native accented speech of the sociolinguistic factor of ethnicity. Rubin shows that perceived ethnicity influences significantly the evaluation of a speaker's language competence as well as his intelligibility and comprehensibility. The former has been discussed, for example, by Major (2007) as well as Kolly and Dellwo (2013). Major shows that, for the recognition of non-native accented speech, the knowledge of L2 is less important than the knowledge of the speaker's L1. Kolly and Dellwo suggest that listeners are able to recognize the speakers' region of origin in Switzerland, even if they speak in L2.

We agree with Major that the L1 is important for the recognition of non-native accent, as well as for the overall perception of non-native accented speech; however, we would rather like to draw attention to the importance of sound similarity between languages. This is mainly viewed through the lens of existing genealogical or typological relations between languages, whereas it should also be bear in mind that sound resemblances may occur even between genealogically or typologically unrelated languages. The topic of sound similarity between non-related languages has first been addressed by

Bradlow, Clopper, Smiljanic, and Walter (2010). In their study Bradlow *et al*. have suggested developing a perceptual phonetic similarity space for languages. They show that certain languages which are not known to be genealogically related may be consistently perceived as being sound-related.

In our paper, we suggest that Lusophones' Czech is constantly perceived by Czech recipients as being of Slavic origin. First, we present the results of our perceptual test which confirms our suggestion. Further, we hypothesize that this perception of Lusophones' speech may have consequences for Czech native speakers' attitude towards speakers of Portuguese and Brazilian origin. This hypothesis is supported by our corpus-based research which implies different connotations/meanings of particular foreign accents in Czech social contexts. We show the different occurrences of differently accented Czech in Czech media, the topics mentioned range from positive descriptions, over completely neutral ones to those connected to perpetrators and victims of a crime. We suggest that the evaluations of different non-native accented speech may be completely diverse depending on whether they are based on linguistic or social cues.

Perception experiment

Speech material

The material consists of two excerpts from a read Czech text:

> *(1) a povídá no tak dobrá tak si jdi a pustil ji ale nějak mu to otrávilo rybolov tak složil pruty a jde domů* [ʔa ˈpoviːdaː ˈno tak ˈdobɾaː tak sɪ ˈjɟɪ ʔa ˈpuscɪl jɪ ˈʔalɛ ˈɲɛjak mu to ˈʔotɾaːvɪlo ˈɾɪbolof tak ˈsloʒɪl ˈpɾutɪ ʔa ˈjdɛ ˈdomuː]

> *(2) a že zkrátka ať jde rybář a řekne rybě že potřebujou vilu vilu na co jen jdi ryba to ráda udělá dala chalupu dá vilu* [ʔa ʒɛ ˈskɾaːtka ʔac ˈjdɛ ˈɾɪbaːr̝̊ ʔa ˈr̝ɛknɛ ˈɾɪbjɛ ʒɛ ˈpotr̝̊ɛbujo͡u ˈvɪlu ˈvɪlu ˈna t͡so jɛn ˈjɟɪ ˈrɪba to ˈraːda ˈʔuɟɛlaː ˈdala ˈxalupu ˈdaː ˈvɪlu]

These sentences were chosen rather randomly, though it was considered important that they would comprise most of the existing Czech phonemes. In total, 26 speakers were selected, out of these 10 are native Portuguese speakers, 10 are native Slavic speakers and 6 are native speakers of non-Slavic languages (excluding Portuguese native speakers). Slavic languages are represented by five Czech native speakers followed by a more or less even number of Russian, Ukrainian, Polish and Slovak native speakers. Each non-Slavic language (besides Portuguese) is also represented by a similar number of speakers; these are native speakers of German, Spanish, English and Japanese. Out of ten Lusophone's, six were native speakers of European Portuguese (EP) and four native speakers of Brazilian Portuguese (BP). From each speaker two speech samples were selected, the test comprised 52 stimuli (2 stimuli from 26 speakers). Duration of one stimulus ranged from 10 to 15 seconds. The authenticity of modified stimuli was verified by native Czech speakers beforehand. According to the Common Reference Levels the competence in Czech of selected speakers ranged from A2 to C1. However, this competence description does not reflect their pronunciation abilities as is clear from available perceptual ratings (see Figure 2 below).

Methodology

Native Czechs listened to the stimuli and performed a classification task based on the assumed L1 of the speakers and ratings of accentedness. For the classification task, the listeners were given four language categories: (a) native Czech, (b) Russian or Ukrainian, (c) other Slavic languages, i.e. Polish or Slovak, etc. and (d) non-Slavic languages. The 'broad' non-Slavic group was selected because our aim was not to test the knowledge of different languages by listeners. The Slavic groups (a, b, c) were selected according to our assumption that Czech native speakers share common knowledge about Russian accent and other Slavic accents, in general. This assumption stemmed from the similarity of segmental features together with differences at suprasegmental level in the Slavic language family. The Czech native speakers are well aware of both, the similarities and the differences thanks to shared historical background and the presence of a lot of foreign residents of Slavic origin in the Czech Republic (see the Figure 1). The awareness of Slavic non-native accent in Czech is also supported by several occurrences of lampooned Russian and Slovak speech. This occurs even in dubbed foreign movies, where accented speech is mimicked by native Czech actors.

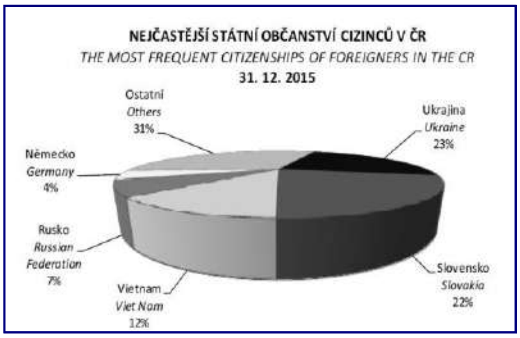


Figure 1. The most frequent citizenships of foreigners in the CR

In the rating task, the listeners were asked to rate the speaker's accentedness, or goodness of one's speech. The accent of speakers was evaluated on a 10-point scale, where 1 stood for a very poor speaker and 10 for an excellent one. The listeners were made aware of the fact that a Czech native speaker did not always have to obtain maximum points possible. It was also allowed to give half a point.

In total, 21 listeners took part in the perceptual test, including four phoneticians. The rest of the listeners had no linguistic or phonetic background. All of the participants were unaware of the aim of the study. The stimuli were presented in randomized order which was different for each listener. In the test, it was possible to listen to each recording repeatedly. The number of replays was not limited, but even then one recording was played at most three times. All activity of the participants while performing the perceptual test was captured by the use of the Google Chrome add-on Cattura: CaptureCast. Thanks to these recordings of the computer screen it is possible to find out where the listeners faced doubts and when they changed their decisions.

Results

In all language groups it is visible that listeners do assign the speakers non-randomly to the correct language group based on speaker's L1. Czech native speakers were thus most often placed into the language group (1) native Czech speakers. It so happened with the other three groups, as well. Table 1 shows the breakdown of the results (in percentages) from all listeners. The group of Portuguese native speakers is a part of group (4) non-Slavic languages. The placement of all speakers into the correct L1 group is based on the results of the chi-square test statistically non-random ($p < 0.0001$).

Table 1. Speaker's placement into 4 language groups (in percentage)

Σ = n recordings / Σ = n judgments	CZ	RU, UA	SK, PL	non-Slavic
CZ	84,3%	1,4%	11,0%	3,3%
RU, UA	0,0%	57,1%	35,7%	7,1%
SK, PL	2,4%	22,6%	54,8%	20,2%
non-Slavic	1,8%	21,0%	21,4%	55,8%

When we further subdivide the non-Slavic group and look on Portuguese speakers, we can again see that they were assigned non-randomly into the correct language group, that is, non-Slavic languages. For additional scoring,we decided to merge all three Slavic groups into one superior Slavic language group. When comparing the results between the non-Slavic and the merged Slavic group we still did not obtain any significant difference. After leaving aside the two worst evaluated speakers the results changed. For those speakers whose language competence was perceived as acceptable (they obtained a minimum of 5 points on the 10-point scale), it is true that they are judged significantly more (p =

0.00011) as being L1 speakers of a Slavic language than as being non-Slavic. Table 2 shows the breakdown of results after excluding the two worst evaluated speakers.

Table 2. Placement of Portuguese native speakers into 4 language groups after excluding 2 worst evaluated speakers (in percentage)

Σ = n recordings / Σ = n judgments	CZ	RU, UA	SK, PL	non-Slavic
PT	3,6%	26,8%	28,6%	41,1%
			Slavic	non-Slavic
			58,9%	41,1%

A comparison of the results obtained from the classification task and those obtained from the evaluation task shows a high correlation (r = 0.754, $p < 0.001$) between evaluation of the goodness of a speaker and the frequency of the speaker's assignment to one of the three Slavic language groups. This high correlation is shown in Figure 2.

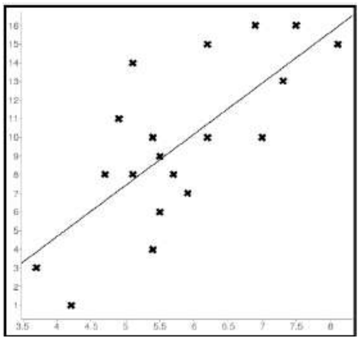


Figure 2. Correlation between "goodness of speaker" (horizontal axis) and assignment to a Slavic L1 group (vertical axis)

A significant variable as the analysis of results suggests seems to be the role of sex. Table 3 shows the breakdown of the results (in percentage) for the assignment of female and male speakers of EP and BP to each language group. The only statistically significant difference ($p < 0.01$) occurs in the assignment to the Russian or Ukrainian group. Other differences in the assignment of Lusophone female and male speakers were not statistically significant ($p > 0.05$).

Table 3. Placement of Portuguese native speakers into language groups based on sex (in percentage)

Σ = n recordings / Σ = n judgments	CZ	RU, UA	SK, PL	non-Slavic
female	0,0%	35,0%	26,3%	38,8%
male	4,2%	15,4%	22,5%	57,9%

Corpus-based study

The perceptual test confirmed that the Lusophones whose Czech were assessed as good were perceived as speakers whose L1 is Russian or other Slavic language. Another step of our research then was to find out whether this fact might have some consequences for the Czechs' language attitudes towards the Lusophones, while speaking Czech. These attitudes could be twofold: a) towards the speakers' language proficiency in L2, and b) towards the speakers' assumed national origin. We are interested in both sides of the attitudes, but because of lack of space only the social aspects of language attitudes are discussed here.

An attitude is defined by Oppenheim (1982:32) as ‘a construct, an abstraction which cannot be directly apprehended. It is an inner component of mental life which expresses itself, directly or indirectly, through such more obvious processes as stereotypes, beliefs, verbal statements or reaction, ideas and opinions, selective recall, anger or satisfaction or some other emotion and in various other aspects of behaviour.’ As an attitude cannot be directly apprehended, we decided to find out which language attitudes Czechs express to the speakers with a foreign accent in Czech by using ‘linguistic worldview’ (Bartmiński, 2009).

Methods

Two corpora of the Czech National Corpus were exploited for gaining such a linguistic worldview. They were, in particular, syn2013pub and syn2015 which are based on journalistic texts from the period 2005-2014. These two specific corpora were chosen because we were mainly interested how the speakers with different non-native accent are presented in Czech press. Beside this, sufficient quantities of examples reside in the chosen corpora. To find the word expressions concerning the non-native accent we used the command (meet[lemma=*ruský*][lemma=*přízvuk*]-4 4). The first lemma (*ruský* ‘Russian’) was changed according to the particular language, the second lemma (*přízvuk* ‘accent’) remained unchanged. The searched language accents were Russian, Ukraine, Slovak, Polish, Spanish, and Portugal. The found collocations were sorted into eight semantic groups: 1. Crime, 2. Movie, 3. Wealth, 4. Politics, 5. Art, 6. Life of foreigners in general, 7. Czech language abroad, and 8. a Czech speaker who mimics non-native accent. Some of these groups were further divided into other small subgroups. This division was needed especially for the first group, Crime. The subgroups were as follows: prostitution, perpetrator of a crime, and a victim of a crime.

Results

Russian accent

The Russian accent was the most frequently found accent of the searched ones. We gained 215 hits in the syn2013pub corpus, but only seven in syn2015. The expression *Russian accent* was mostly mentioned in relation to crime (49%); 23% of its occurrences was related to the Russians’ life in general, and 14% to Russian-accented speech in movies. Other topics concerning Russian accent were discussed only sporadically (14%). Figure 3 shows the breakdown of the results.

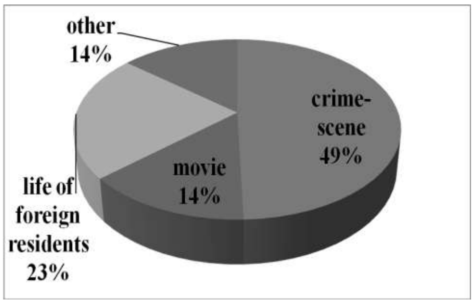


Figure 3. Russian accent in SYN2013pub corpora

A more interesting picture of Russian accent as represented in newspapers can be found with a detailed look on the subgroups of the most frequented topic criminality. Such subdivision shows that Russian accent tends to be mostly connected to commitment of crime, this is the case of overall 97%. No occurrence showing a Russian accented person as a victim of a crime was found. The breakdown of the results is shown in Figure 4.

Slovak accent

The Slovak accent was significantly less frequent than the Russian one. We gained 90 hits in syn2013pub corpus and only one hit in syn2015. Similarly to the Russian accent, the connection with crime predominates also in relation to Slovak accent (63%). The breakdown into all semantic groups is shown in Figure 5. The only subgroup found in connection with Slovak accent is the perpetrator of a crime.

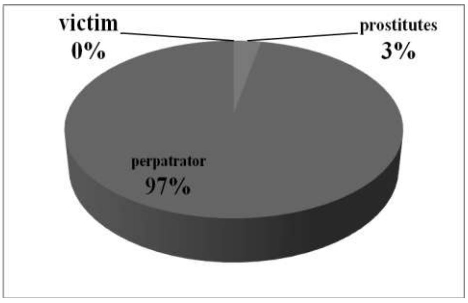


Figure 4. Different types of 'role' on crime-scene (Russian accent)

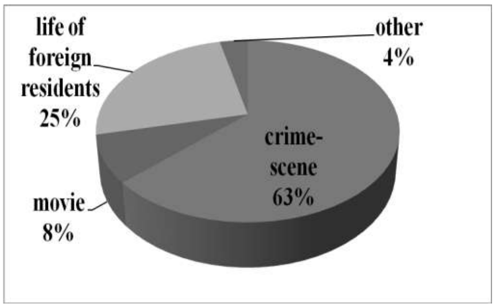


Figure 5. Slovak accent in SYN2013pub corpora

The following excerpt from an article found in the corpus shows how Czech media are used to dealing with different Slavic accents. The perpetrator of a crime is described here as a person who spoke Czech with either a Slovak or Russian accent. This 'underdifferentiation' of Slavic accents in Czech is frequently found in Czech media. An important category used to describe foreign accents in Czech is the so-called 'eastern' accent. This category seems to be used for describing the accent from any person who might come from a post Soviet country.

> Právo, 8.12.2005 (syn2013pub)
> [...] zaměstnankyni k otevření trezoru a vydání finanční hotovosti , " uvedl mluvčí táborské policie Jaroslav Doubek . Jednalo se o muže ve věku mezi 20 a 25 lety , střední postavy . Na sobě měl světle šedou zimní bundu , modré riflové kalhoty , na hlavě šedou pletenou čepici . Mluvil česky s mírně slovenským či ruským přízvukem . Pálil dívku cigaretou [...]
>
> [...] the employee to open the safe deposit and to issue cash, stated the spokesman of Tábor policy, Jaroslav Doubek. He was a man between 20 and 25 years of age, of average height. He was wearing a light grey winter jacket, blue jeans, and grey knitted cap. He spoke Czech with a slight Slovak or Russian accent. He was burning a girl with a cigarette [...]

Polish accent

Slightly different are connotations of the Polish accent. This is connected to crime only at 9% and, most frequently, it is described in relation to neutral topics regarding the life of foreign residents in the Czech Republic.

Spanish and Portuguese accent

The Spanish and Portuguese accents are represented only with 16 hits in the corpus, while the Portuguese accent is never discussed in relation to the Czech language. The Spanish accent is connected to crime only in one occurrence; otherwise, it is discussed in relation to Antonio Banderas and Osmany Laffita (a fashion designer originally from Cuba, who lives in the CR and speaks Czech fluently). As opposed to the Russian and Slovak accents, the Spanish accent is pictured with positive attributes, such as 'the Czech with a nice Spanish accent' or 'the cute Czech with the Spanish accent'.

Discussion and conclusion

The participants in the perceptual test had two tasks: to sort the speakers according to their assumed L1 into one of four language group; to assess how good or bad was their accent when they read the Czech fictional text. When we correlate results of these two tasks, we can see a rather clear picture concerning not only the Czech native speakers, but also the Lusophone ones. Not surprisingly, highly assessed speakers were marked as Czech native speakers. But with regard to the Lusopohone speakers, the results were not so predictable. The Lusophone speakers to whom 5 points or more were assigned were mostly put into the Russian and Ukrainian language group, or in some cases, into another Slavic language group. Only those Lusophone speakers whose Czech was evaluated with less than five points were placed into the non-Slavic language group. In other words, when a Lusophone speaker could master reading the Czech fictional text very well, s/he was perceived as a native speaker of Russian or other Slavic language. This will be a topic of our future research. Right now, however, we can state that the fact of being mostly perceived as speakers with a Slavic L1 has at least two consequences for the Lusophones speaking Czech. The language attitudes are built in fact on the assumption of the particular L1 and on the assumption of the particular social background, as well. The assumption of the Czechs that the L1 of a foreigner is Russian or other Slavic language increases expectancy about their ability to speak good Czech. And the assumption of Slavic origin of speakers carries preconceptions and stereotypes about them. Particularly, speakers with a Russian accent are conceived by the Czechs as suspicious people. These speakers are often perceived in the Czech Republic as people associated with crime, or with doubtfully gained wealth. They can be also associated with workers from Ukraine who work on construction sites, usually doing inferior work.

It is also remarkable that female Lusophone speakers are more often than their male counterparts perceived as speakers whose L1 is Russian. Why Lusophone women are more often than Lusophone men perceived as native speakers of Russian will be also the subject of further research. It would also be appropriate to focus on the particular segmental and suprasegmental features that influence the perceptual assessment of Lusophones' non-native accent by Czechs.

References

Bartmiński, J. (2009). Viewpoint, perspective, and linguistic worldview. in: J. Zinken (ed): *Aspects of Cognitive Ethnolinguistics* (pp. 76-87). London, Oakvill: Equinox.

Bradlow, A., Clopper, C., Smiljanic, R., & Walter, M. A. (2010). A perceptual phonetic similarity space for languages: Evidence from five native language listener groups. *Speech Communication*, *52*(11-12), 930-942.

Kolly, M.-J., & Dellwo, V (2013). (How) can listeners identify the L1 in foreign accented L2 speech?. *Travaux neuchâtelois de linguistique*, *59*, 127-148.

Křen, M., Cvrček, V., Čapka, T., Čermáková, A., Hnátková, M., Chlumská, L., Jelínek, T., Kovářiková, D., Petkevič, V., Procházka, P., Skoumalová, H., Škrabal, M., Truneček, P., Ondřička, P., & Zasina, A. (2015). *SYN2015: Reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK. Online: http://www.korpus.cz

Křen, M., Hnátková, M., Jelínek, T., Petkevič, V., Procházka, P., & Skoumalová, H. (2013). *SYN2013PUB: korpus psané publicistiky*. Praha: Ústav Českého národního korpusu FF UK. Online: http://www.korpus.cz

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, *46*(6), 1093-1096.

Major, R. C. (2007). Identifying a foreign accent in an unfamiliar language. *Studies in Second Language Acquisition*, *29*(4), 539-556.

Oppenheim, B. (1982). A exercise in attitude measurement. In G.M. Breakwell, H. Foot & R. Gilmour (eds.), *Social Psychology: A Practical Manual* (pp. 38-52). London, UK: Macmillan Press.

Rubin, D. L. (1992). Non-language factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*(4), 511-531.

To cite this article: Zaepernickova, E., Havlik, & M. Kraftová (2017). Lusophones’ pronunciation of Czech and its perception by native Czech speakers. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 301-307). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

Swedish for immigrants. Teachers' opinions on the teaching of pronunciation

Elisabeth Zetterholm
elisabeth.zetterholm@isd.su.se

Stockholm University

Abstract. This study examines teachers' opinions on the teaching of pronunciation in Swedish for immigrants (SFI). A web-based survey containing questions about teachers' experiences of teaching pronunciation and their education in second language pronunciation was distributed to teachers in Sweden. The purpose of the survey was to learn more about how pronunciation is taught to adult learners of Swedish as a second language. 92 teachers of SFI answered questions about their approach and beliefs concerning teaching pronunciation. Almost all of them reported that it is difficult to teach pronunciation because they do not have enough training in teaching it. Only half of the teachers have any kind of formal education in pronunciation methodology. However, they all consider pronunciation instruction to be of importance for second language learners. The goal for second language teaching in general is to obtain comprehensible speech in order to have an increased chance of integration into the society. Only a few teachers have special lessons focusing on pronunciation. Most of them take the opportunity to talk about and teach pronunciation embedded in lessons concerning, for example, grammar. The teachers answer that vowel quality, phrase intonation and word stress is of importance when teaching pronunciation, but they also mention that prosody is difficult to teach. The results of this survey raise questions regarding teachers' education in teaching pronunciation and the teachers' knowledge about Swedish phonology and phonetics.

Keywords: pronunciation teaching, teachers' education, second language teaching

Introduction

For both learners and teachers of a second language, pronunciation holds importance. For many adult learners, a foreign accent in the second language is an unavoidable eventuality and one potential marker of the speaker's identity (e.g., Moyer, 2013). The degree of foreign accent can be crucial in how the speaker is perceived by listeners, though the connection between accent and intelligibility is not clear (Munro & Derwing, 1995). So, while individual learners may strive for native-like pronunciation in the second language, the goal for all learners in the classroom should instead be to achieve intelligible pronunciation due to its importance for everyday communication and integration into the society.

With this in mind, it is important for teachers to be equipped with the knowledge of how to teach pronunciation and to include it in their teaching. Studies have shown that pronunciation instructions lead to improvement and that learners' perception and production can improve as a result of teaching instructions (e.g., Derwing & Munro, 2015; Thomson & Derwing, 2015) However, there appears to be an underrepresentation of pronunciation instructions in second language teaching, because teachers' do not receive training on it in their teacher education (e.g., Breitkreutz, Derwing, & Rossister 2001; Foote, Holtby, & Derwing 2011). It is fundamental that pronunciation methodology be taught at university courses for teachers-to-be. Several studies conducted in English-speaking countries on pronunciation teaching conclude that guidelines for teacher training are not well defined (e.g., Burns, 2006; Burgess & Spencer, 2000; Darcy, Ewert, & Lidster 2012; Deng, Holtby, Howden-Weaver, Nessim, Nicholas, Nickle, Pannekoek, Stephan, & Sun, 2009; MacDonald, 2002). The results presented in Foote *et al*. (2011) indicate that teachers in their study indeed understand the importance of pronunciation instructions but express a desire for more pronunciation training in their teacher education. In terms of Swedish as a second language, little research has been carried out on how and to what extent teachers of adult learners teach pronunciation, on their opinions surrounding pronunciation instructions and their own teacher training, and so they deserve further attention.

The aim of the study

The aim of this study is to learn more about of how pronunciation is taught to adult learners of Swedish as a second language. It aims to answer the following questions: (1) what is their opinion regarding the importance of pronunciation instructions?, (2) how much do they teach pronunciation and how?, (3) do Swedish-as-a-second-language teachers receive any pedagogical training in pronunciation at university?, and (4) are they prepared to teach pronunciation?

Method and material

The answers from a web-based survey containing questions on teachers' experiences of teaching pronunciation and their training in second language pronunciation are analysed in this study. The questions centred in on whether, what, how and when they teach pronunciation. The survey included yes/no, multiple choice, as well as open-ended questions. For some questions participants were given the option to write an additional comment voluntarily, but this was not compulsory.

Participants

92 teachers in the program Swedish for Immigrants (SFI) participated in this study, the majority of whom were female (88%). While there were a variety of other European first languages among the participants, the majority regard Swedish as their first language (80/92). Six participants were less than 30 years old and 11 older than 60 years, but most of the participants were between 30-60 years old (see Figure 1).

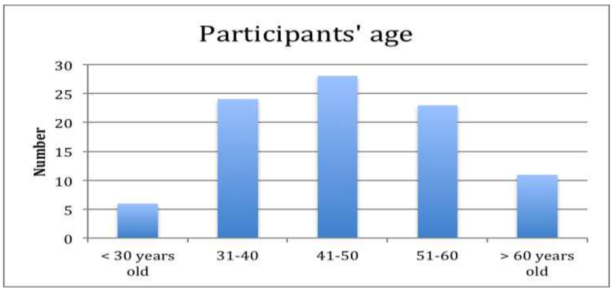


Figure 1. The age of all participants

To be a certified teacher in Swedish as a second language, one is required to have studied one semester at a university, during which time he or she has to pass four courses. The content of the courses often differs across universities. However, many students take extra courses since they consider one semester to be insufficient for teaching adult learners of Swedish as a second language. In this study, 91% of the participants answer that they are certified teachers and as many as 76% have studied two or more semesters at a university, though most of them two semesters (see Figure 2). Nevertheless, only 46% of the participants say that pronunciation instructions and methodology were a part of their education.

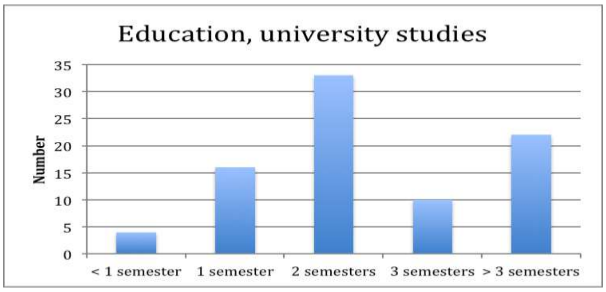


Figure 2. Participants' education

Results

In response to questions regarding the teachers' overall goal when teaching Swedish as a second language, many consider the most important goal to be that learners can read and understand different kinds of texts, different genres in Swedish as well as grammar. Most frequent answers were (in ranked order):

- Read and understand different texts in Swedish, different genre and fictions
- Grammar
- Reading and writing in Swedish
- Vocabulary
- Communication
- Conversation in everyday life

Some of the teachers also mention the four goals in the syllabus for SFI: listening comprehension, reading comprehension, oral interaction, writing skills. In teachers' comments it is clear that most of them think that communication is of importance which indicates that pronunciation teaching should also be of importance. When asked more specifically about the goals of pronunciation teaching, the participants' responses include:

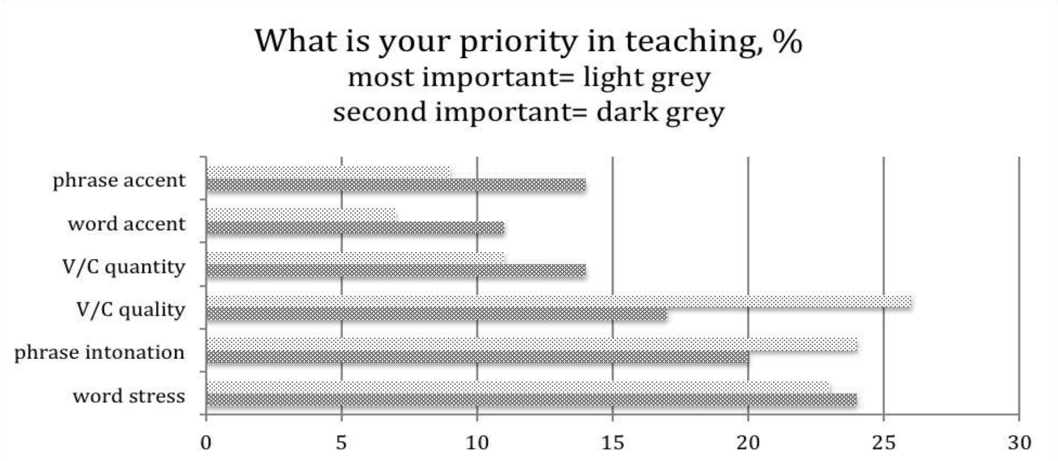
- Achieve an intelligible speech in Swedish
- Achieve a Swedish-like pronunciation
- Listener-friendly pronunciation
- As close to native-like pronunciation as possible

When you teach pronunciation – is it explicit or implicit?

Only 12% of the teachers answer that they have pronunciation lessons where they talk about Swedish pronunciation explicitly, sometimes using digital resources. The most common way is to listen to a recording of a native speaker or the teacher and to imitate the speech. This is often performed chorally, initiated by the teacher using instructions, such as one participant who writes *repeat after me, please*. 44% of the teachers include pronunciation instructions in other lessons where the main focus is on grammar or word acquisition. Some teachers (44%) answer that they talk about pronunciation when teaching reading and writing in Swedish. While not all respondents left written comments, quite a few of the participants that did so, say that they often talk about pronunciation and correct the learners' pronunciation. Maybe only those who really teach pronunciation left written responses as a complement and clarification to their answers. Only a few teachers say that listening is of importance and mention that they use poetry both for listening and reading, with a focus on pronunciation. Two participants say that they sing together in the classroom. When analysing all the answers, it is clear that many teachers think pronunciation is of importance and take the opportunity to talk about Swedish pronunciation on different occasions, but they do not include many explicit pronunciation instructions, such as articulation. Another perspective of teaching pronunciation is whether the teacher can explain differences between the learners' first language and the target language (here Swedish) in a contrastive way, which might help some learners to be aware of both articulatory and phonological differences as well as similarities.

What is your priority when teaching pronunciation?

This was a multiple choice question. Comments from teachers indicate that they did not fully understand the question and the differences between some of the suggestions, and so they did not know how to answer. Some of them say that everything is equally important and it is not possible to focus more on, for example, intonation or word stress. However, the answers overall indicate that teachers often think that the quality of vowels and consonants is very important in Swedish, followed by phrase intonation and word stress. In Swedish, vowel quality is connected to and contributes to vowel quantity, which is quite important for an intelligible speech in Swedish. A stressed syllable in Swedish has longer duration and a slightly higher pitch compared to an unstressed syllable; this is so, not only with regard to the vowel in the syllable but also the consonant, and this quantity distinction is complementary. The syllable structure in Swedish is a long vowel followed by a short consonant or a short vowel followed by a long consonant (V:C or VC:) (e.g., Zetterholm & Tronnier, 2017). Word stress is a distinctive feature in Swedish, and since there are several minimal pairs with that distinction, it is important to be aware of stress and the quantity of the syllable. The teachers' choices of the most and less important features in teaching pronunciation are shown in Figure 3.



	word stress	phrase intonation	V/C quality	V/C quantity	word accent	phrase accent
Serie1	23	24	26	11	7	9
Serie2	24	20	17	14	11	14

Figure 3. Teachers' priority when teaching pronunciation

What kind of education concerning pronunciation teaching have you got?

Almost half of the participants (46%) answer that they have had some kind of education concerning pronunciation teaching. Some mention having taken some university courses, while others mention further education for teachers. Many of the participants express a desire for more education with a specific focus on pronunciation methodology and textbooks with clear instructions. Comments such as *I really want to know more about how to teach pronunciation* indicate that we have to look into the content of the teacher training courses for teachers in Swedish as a second language.

Discussion

Teachers of Swedish as a second language for adult learners consider reading, writing and grammar to be the most important for second language learners of Swedish. Oral skills are not the primary goal according to the spontaneous answers in this survey, taken by 92 teachers in different places all over Sweden. However, it is clear that teachers think that pronunciation is of importance for intelligible speech and for communication in everyday life. A native-like pronunciation is mentioned to be a goal even though listener-friendly and comprehensible pronunciation is more attainable by most adult second language learners. The results indicate that teachers have not received enough education in pronunciation instruction and methodology which is why they do no teach pronunciation explicitly, and they are not really prepared for that. More than 50% of the teachers say that they are not prepared for pronunciation teaching, because they haven't been adequately trained in pronunciation methodology. Comments and answers about the teachers' priority when teaching pronunciation indicate that there is lack of knowledge of Swedish phonology, especially concerning the connection between quantity and word stress. None of the teachers mentioned taking pronunciation lessons with instruction on articulation and contrastive comparisons with the learners' first language. For some learners, clear instruction and comparison between languages might lead to better understanding of Swedish pronunciation and of how to achieve listener-friendly pronunciation. Greater differences might be easier to perceive and produce, but very subtle differences might need further explanation.

The results of this survey are in accordance with other studies on pronunciation teaching and the lack of instruction courses and pronunciation methodology in teacher education (e.g., Foote *et. al*., 2011).

Conclusion

Teachers' opinions about teaching pronunciation and questions surrounding their own teaching have been raised in this study. Results indicate that teachers are aware of the importance of pronunciation teaching, but as many as half of the participants mention that they have not been trained how to teach pronunciation and, therefore, they are not confident in teaching it. They want more pronunciation instruction and methodology in their education. The content in university courses for teachers-to-be merits further investigation in light of the participants' answers in this study, since pronunciation is crucial for second language learners' communication in everyday life.

References

Breitkreutz, J. A., Derwing, T. M., & Rossiter, M. J. (2001). Pronunciation Teaching Practices in Canada. *TESL Canada Journal, 19*(1), 51-61.

Burgess, J., & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System*, *28*, 191-215.

Burns, A. (2006). Integrating research and professional development on pronunciation teaching in a national adult ESL-program. *TESL Reporter*, *39*, 34-41.

Darcy, I., Ewert, D., & Lidster, R (2012). Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation 'toolbox'. In J Levis & K. LeVelle (eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference* (pp. 98-108). Ames, IA: Iowa State University.

Deng, J., Holtby, A., Howden-Weaver, L., Nessim, L., Nicholas, B., Nickle, K., Pannekoek, C., Stephan, S., & Sun, M. (2009). *English pronunciation research: The neglected orphan of second language acquisition studies?* (WP 05-09). Edmonton, AB: Prairie Metropolis Centre.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals. Evidence-based perspectives for L2 teaching and research*. Amsterdam/Philadelphia: John Benjamins.

Foote, J. F., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada 2010. *TESL. Canada Journal, 29*(1).

Moyer, A. (2013). *Foreign accent. The phenomenon of non-native speech*. Cambridge, UK: Cambridge University Press.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73-97.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics, 36*(3), 326-344.

Zetterholm, E., & Tronnier, M. (2017). *Perspektiv på svenskt uttal. Fonologi, brytning och didaktik*. Lund: Studentlitteratur.

To cite this article: Zetterholm, E. (2017). Swedish for immigrants. Teacher's opinions on the teachind of pronunciation. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 308-312). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

The effect of argument omission on monolingual and bilingual children: Evidence from the acquisition of Japanese case-marking

Akiko Zhao Chou
zcza88@gmail.com

Hiroshima University

Abstract. This study investigates how monolingual and bilingual children recognize speech which uses argument omission. Specifically, this study focuses on object case-markings in Japanese transitive sentences where argument omission occurs at a high percentage. In study 1, the effectiveness of argument-omitted sentences and full-argument sentences on monolingual children's learning (ages four to seven) of artificial case-markings via short exposure was compared. Meanwhile, the ability of bilingual children (ages four to nine) to learn artificial case-marking through exposure to argument-omitted and full-argument sentences was investigated in study 2. In study 1, the monolinguals first watched and imitated four single action scenarios while listening to sentences with two non-lexical syllables, *po* (artificial subject marker) and *bi* (artificial object marker). Half of the participants learned full-argument sentences, and the rest learned argument-omitted sentences. In another test, participants completed forced-choice discrimination of scenarios after hearing sentences with either *po* or *bi*. Contrary to study 1, bilinguals in study 2 were only subjected to argument-omitted sentences. A mixed effect model for children's responses (correct or incorrect answer) showed that in study 1, children who learned case-markings through argument-omitted sentences comprehended OSV sentences well. However, children who learned case-markings through full-argument sentences did not do well in sentence comprehension. Results indicated that argument-dropped exposure, not full-argument exposure, was useful for monolinguals in learning object case-markings. For study 2, children who learned case-markings through argument-omitted sentences also comprehended OSV sentences well as compared to chance level. Results showed that both monolinguals and bilinguals probably pay attention to case-markings when they process argument-dropped sentences since this is the only way to understand a sentence. This study suggests that simpler sentences may work better for learning purposes, specifically for languages with high percentages of argument omission.

Keywords: case-marking, argument omission, linguistic information, monolinguals, bilinguals

Introduction

Decades of research have documented argument information such as nouns, verbs, adjectives and prepositions as a useful cue for children to learn words (e.g., Arunashalam & Waxman, 2010; Mintz & Gleitman, 2002; Fisher, Klingler, & Song, 2006; Waxman, Lidz, Braun, & Lavin, 2009). For example, experimental studies of word learning have demonstrated that two year olds utilize argument information to distinguish whether a word is a noun or a verb (Arunachalam & Waxman, 2010). In the experiment, children listened to two types of sentences varied by condition: non-syntactic information condition ('Let's see a boy and a balloon. Let's see pilking') and syntactic information condition ('A boy is gonna pilk a balloon. Let's see'). Then, children saw these scenes side-by-side. In one scene, the boy was performing a familiar action, but on a new object (e.g., waving a rake); in the other scene, the boy was performing a new action, but with a familiar object (e.g., tapping the balloon). The experimenter then asked the children to choose one scene that matched the sentence that they listened to. The result showed that the participants successfully mapped the verb 'pilk' to the action. This result suggests that two year olds could utilize syntactic information to distinguish whether a word is a noun or a verb (see also Imai, Haryu, Okada, Lianjing, & Shigematsu, 2006).

However, in several languages, such as Chinese, Korean and Japanese, both adults and children tend to drop arguments in their utterances. For instance, case-markings are often omitted in Japanese sentences (Matsuo, Kita, Shinya, Wood, & Naigles, 2012; Omaki, Kobayashi, Lassotta, Rizzi, & Franck, 2012). According to Matsuo *et al*. (2012), parent's utterances with transitive verbs appeared with both overt subjects and objects just over 1.3% (e.g., *kuma-ga usagi-wo oshita* ‘bear-NOM rabbit-ACC push-past’), and objects appeared with *wo* over 5% of the time (e.g., *usagi-wo oshita* ‘rabbit-ACC pushe-past’). Thus, Japanese children have very few opportunities to hear sentences with argument information (i.e. full-argument sentence). Arunachalam, Leddon, Song, Lee, & Waxman (2013) also investigated the effects of argument omission in language learning. Specifically, they focused on how Korean children learn verbs. In Korean[1], it is difficult for children to utilize argument information to learn verbs, since they have very few opportunities to hear sentences with argument information (i.e. full-argument sentence). Concretely, children learning Korean have difficulty learning the meanings of novel transitive verbs, if they appear in a full-argument sentence with two arguments (e.g., subject and object), than in an argument-omitted sentence with both arguments elided (Arunachalam *et al*., 2013). The goal of this study therefore was to examine how argument omission affects Japanese object case-marking learning.

Japanese monolingual children start to use case-marking when they are two years old. However, in this period, they do not know the meaning of case-marking and they also cannot utilize it to identify subjects and/or objects in sentences. 3 to 4-year-olds acquire the knowledge of case-marking *ga*, and five to six year olds acquire *wo*. By the time children turn five or six years old, they start to identify subjects and objects using their case-marker knowledge. Before acquiring the object case-marker *wo*, both monolingual and bilingual children utilize word order to identify subject and object in the sentence. For example, when children hear *usagi-ga kuma-wo oshita* ‘rabbit-NOM bear-ACC push-past’, children will think *usagi* is the subject, and *kuma* is the object. Because of word order knowledge, children will misunderstand scrambled sentences likes *kuma-ga usagi-wo oshita* ‘bear-NOM rabbit-ACC push-past’ (Hayashibe, 1975; Iwadate, 1980; Suzuki, 2007; Zhao & Sakai, 2014). If children also utilize word order to distinguish case-marker functions, full-argument sentences will be useful for case-marking learning. However, as mentioned above, children have very few opportunities to hear full-argument sentences since case-markings are often omitted in Japanese.

In the current study, we tested the effect of argument omission. Specifically, we investigated if full-argument sentences or argument-omitted sentences are effective for monolingual (Experiment 1) and bilingual children (Experiment 2) to learn object case-markers. We taught artificial subject and object case-markers to participants using full-argument sentences (full-argument condition) and argument-omitted sentences (argument-omitted condition). Then, participants completed forced-choice discrimination of scenarios after hearing OSV sentences with either *po* or *bi*. Our prediction was that, if children utilized word order to learn case-markers, children who learned case-markers in full-argument sentences would comprehend case-markers better than children who learned case-markers in argument-omitted sentences.

Experiment 1

Method

Participants. Twenty-eight 4-year-old (M = 4;6, range = 4;3–4;9), thirty-three 5-year-olds (M = 5;6, range = 5;1–5;11), twenty-eight 6-year-old (M = 6;7, range = 6;3–6;11) and twenty-two 7-year-old (M = 7;5, range = 7;2–7;11) Japanese monolingual children in several nursery schools in Hiroshima or Chiba, Japan participated in this experiment.

Stimuli. For auditory stimuli, we used sentences where two non-lexical syllables were presented, *po* and *bi* (artificial case-marking), which referred to subject and object, respectively. For the sentences, we used full-argument sentence (SOV: *saru-po ushi-bi oshiteiruyo* ‘monkey-NOM cattle-ACC pushing’) and argument-omitted sentence (SV: *saru-po oshiteiruyo* ‘monkey-NOM pushing’, OV: *ushi-bi oshiteiruyo* ‘cow-ACC pushing’) to teach the case-marking to the participants.

For visual stimuli, animal moving videos were presented during both the learning and test phases. The animals used were the following: monkey, cow, rabbit, bear, elephant, tiger, panda and lion. For instance, in the learning phase, a video showing a monkey pushing a cow was presented. In the test

[1] Japanese child-directed speech has similar rates of argument omission as Korean (Y.-J. Kim 2000).

phase different animal pairs (e.g., rabbit-bear, elephant-tiger and panda-lion) with different actions (e.g., kick and pull) were presented (Table 1).

Procedure. We taught children the artificial case-markings *po* and *bi* in two types of argument structures: (a) full-argument sentence that can utilize the word order and (b) argument-omitted sentence that cannot utilize word order. In the experiment, participants initially completed forced-choice discrimination of scenarios by listening to eleven sentences (Eight O-*bi*-S-*po*-V sentences as a pre-test, and one sentence for each of these as a filler sentence: S-*po*-V, O-*bi*-V, S-*po*-O-*bi*-V). Then, they watched and imitated single-action-scenarios while listening to sentences where two non-lexical syllables were presented, *po* and *bi*, which referred to agent and patient, respectively. Half of the population sample learned full-argument sentences, and the other half learned argument-omitted sentences. After this, participants completed forced-choice discrimination of scenarios by listening to eleven sentences as a post-test (The test was same as pre-test) (see Table 1).

Table 1. Representative set of stimuli

	Pretest	Learning	Posttest
Audio stimulates	OSV: Laionsan bi pandasan po ketteruyo! lion-ACC panda-NOM kick-prog (The panda is kicking the lion)	**Full-argument sentence (SOV):** *sarusan po ushisan bi oshiteruyo!* monkey-NOM cow-ACC push-prog The monkey is pushing the cow! **Argument-omitted sentence (SV/OV):** SV: *sarusan po oshiteruyo!* monkey-NOM push-prog The monkey is pushing! OV: *ushisan bi oshiteruyo!* cow-ACC push-prog The monkey is pushing!	OSV: Laionsan bi pandasan po ketteruyo! lion-ACC panda-NOM kick-prog (The panda is kicking the lion)
Video Clip			

Results

Correct and incorrect choices were analyzed group by group using Generalized Logistic Mixed Effects Models (GLME). The fixed factors are learning condition and age while the random factors are participants and items. The results were shown in Figure 1 (a, b).

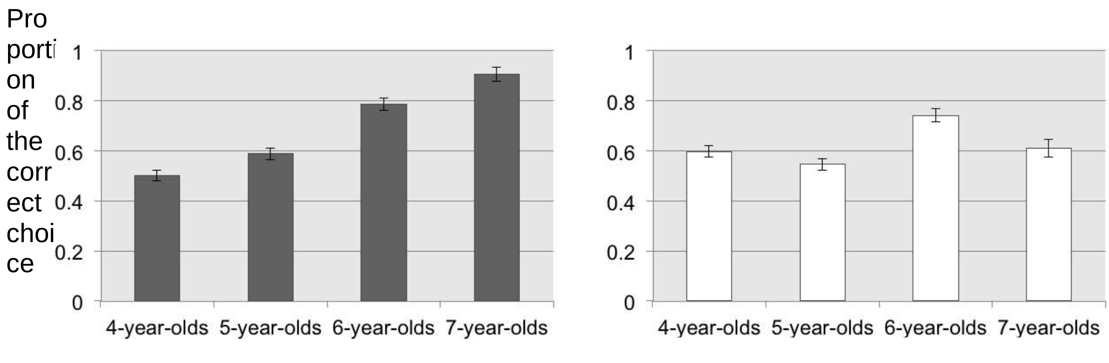


Figure1. Mean proportion of points to the correct scene (a. argument-omitted condition, b. full-argument condition)

There was an interaction between learning condition and age: β= 0.72, SE = 0.22, z = 3.22, p = .001. Post-hoc analysis reviewed that there was a significant difference for ages in the argument-omitted condition (β = 0.91, SE = 0.18, z = 4.99, p < .0001) but not in the full-argument condition (β= 0.19, SE = 0.18, z = 1.06, p = .288). The results above indicate that argument-omitted sentences are more effective than full-argument sentences for Japanese monolingual children in learning object case-markings. In addition, the effectiveness has a developmental difference; specifically, it is effective for the 6 and 7-year-olds.

Experiment 2

Method

Participants. The participants were sequential bilinguals - nine children aged 3 to 9 years old who started to learn Japanese. The profiles are shown in Table 2.

Table 2. The profiles of the bilingual participants

Participant	1	2	3	4	5	6	7	8	9
Age	6	10	10	7	7	8	9	8	7
Exposure	3 mos	3 mos	3 mos	6 mos	7 mos	10 mos	11 mos	36 mos	36 mos
L1	CHI	CHI	CHI	INDO	CHI	KYRG	VIT	NEP	CHI

Stimuli and Procedures. For the artificial case-marking learning experiment, both stimuli and procedures were similar to Experiment 1. The only difference is that only argument-omitted sentences (SV: *saru-po oshita* 'monkey-NOM pushing', OV: *ushi-bi oshita* 'cow-ACC pushing') were used in Experiment 2. Aside from the artificial case-marking learning experiment, the existing case-markings were also tested. For auditory stimuli, we used sentences which used existing Japanese case-markings *ga* (subject marker) and *wo* (object marker). *Ga* and *wo* were presented in both canonical word order sentences (SOV: *onesan-ga onisan-wo oshiteiruyo* 'the girl-NOM cattle-ACC pushing') and scrambling sentences (OSV: *onesan-wo onisan-ga oshiteiruyo* 'the girl-ACC cattle-NOM pushing') (Figure 2). After listening to these sentences, the participants were required to choose a picture which semantically matches them.

Test sentences

SOV sentence : onesan-ga onisan-wo hippateiruyo
girl-NOM boy-ACC push-prog 'the girl is pushing the boy'

OSV sentence: onesan-wo onisan-ga hippateiruyo
girl-ACC boy-NOM push-prog 'the boy is pushing the girl'

Figure 2. Representative set of stimuli

Results

For the artificial case-marking learning experiment, 59 choices (82%) were correct, 13 choices (18%) were incorrect. Chi-square analyses were performed on the number of correct choices for the post-test to compare performance to chance. Performance was significantly above chance ($\chi^2 = 16.364$, $df = 1$, $p < .0001$). The results indicated that the participants were able to learn the object case-marking using the argument-omitted sentence.

For the existing case-marking experiment, the results are as follows. For SOV sentences, 54 choices (100%) were correct and 0 choice was incorrect, while for OSV sentences, 18 choices (33%) were correct and 82 choices (67%) were incorrect. Chi-square analyses for the correct choice revealed that there was a significant difference between the correct choices of the SOV sentences and the OSV sentences ($\chi^2 = 54$, $df = 1$, $p < .0001$). It suggested that participants do not understand the function of object case-marking but they can understand sentences using their word order knowledge.

Discussion and conclusions

The present study examined how argument omission affects object case-marker learning. The result suggested that in learning object case-markings, argument-omitted sentences are more effective than full-argument sentences for both monolingual and bilingual children.

Why are argument-omitted sentences more effective in learning object case-marking than full-argument sentences? This is probably because argument-omitted sentences allow children to pay attention to the target case-marking, and then utilize it to infer the meaning of the sentence. On the other hand, for full-argument sentences, paying attention to case-markings is not necessary for understanding sentences. For instance, if children hear this full-argument sentence: *panda-po lion-bi ketteiruyo* ‘panda-CASE-MARKING lion-CASE-MARKING kick-PROG’, they can infer the sentence meaning using their word order knowledge, not their case-marking knowledge. However, if children hear this argument-omitted sentence: *lion-bi ketteiruyo* ‘lion-CASE-MARKING kick-PROG’, to understand this sentence, they need to pay attention to the case-marking *bi*, and then relate the sentence to the event.

Previous studies revealed that argument information is effective for children to learn new words (e.g., Arunashalam & Waxman, 2010; Fisher *et al*., 2006; Mintz & Gleitman, 2002; Waxman *et al*., 2009). On the other hand, Arunachalam *et al*. (2013) and Imai *et al*. (2006) focused on languages (e.g., Korean, Japanese) that allow a high percentage of argument omission. These studies documented that argument-omitted sentences are effective for children in learning new words, since argument-omitted sentences appear in the input often. The effectiveness of argument-omitted sentence was also confirmed by the present study. Moreover, the present study also made clear that argument-omitted sentences are not only effective for monolingual’s first language acquisition, but also for bilinguals’ second language acquisition.

References

Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition, 114*(3), 442-446.

Arunachalam, S., Leddon, E. M., Song, H. J., Lee, Y., & Waxman, S. R. (2013). Doing more with less: Verb learning in Korean-acquiring 24-month-olds. *Language Acquisition*, *20*(4), 292-304.

Fisher, C., Klingler, S. L., & Song, H. J. (2006). What does syntax say about space? 2-year-olds use sentence structure to learn new prepositions. *Cognition*, *101*(1), 19-29.

Hayashibe, H. (1975). Word order and particles: A developmental study in Japanese. *Descriptive and Applied Linguistics*, *8*(1), 18.

Imai, M., Haryu, E., Okada, H., Lianjing, L., & Shigematsu, J. (2006). Revisiting the noun-verb debate: A cross-linguistic comparison of novel noun and verb learning in English Japanese and Chinese speaking children. In K. Hirsh-Pasek & R. Golinkoff (eds.), *Action meets word: How children learn verbs* (pp. 450-476). New York: Oxford University Press.

Matsuo, A., Kita, S., Shinya, Y., Wood, G., & Naigles, L. (2012). Japanese two-year-olds use morphosyntax to learn novel verb meanings. *Journal of Child Language*, *39*, 637-663.

Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition*, *84*(3), 267-293.

Omaki, A., Lassotta, R., Kobayashi, T., Rizzi, L., & Franck, J. (2012). Delay of word order development in Japanese: Evidence from a preferential looking study with 19- and 30-month-old children. *Poster presented at the 34th Meeting of Cognitive Science Society*.

Suzuki, T. (2007). The development of Japanese case-markers observed through children’s comprehension of single-argument sentences. *Gengo kennkyu*, *121*(132), 55-7.

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants’ interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, *59*(1), 67-95.

Zhao, A. & Sakai, H. (2014). Japanese children's learning of syntactic markers, -*ga* and -*wo*: A sentence comprehension experiment using familiar verb. *JCSS Japanese Cognitive Science Society,* 108-117.

To cite this article: Zhao Chou, A. (2017). The effect of argument omission on monolingual and bilingual children: Evidence from the acquisition of Japanese case-marking. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 313-317). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017

The dynamics of marked consonant clusters in Polish

Paulina Zydorowicz, Katarzyna Dziubalska-Kołaczyk
zpaula@wa.amu.edu.pl, dkasia@wa.amu.edu.pl

Adam Mickiewicz University, Poznań

Abstract. Phonotactics investigates permissible sound combinations in a language. Polish allows for as many as 4 consonants word-initially, 6 consonants word-medially and 5 consonants word-finally. Polish clusters may be phonologically or morphologically motivated. The former are referred to as lexical or phonotactic clusters, whereas the latter are labeled as morphonotactic, meaning that they emerge as a result of concatenative or nonconcatentive morphology. Consonant clusters may have varying degrees of complexity (expressed by markedness or naturalness). The goal of this paper is the analysis of the most marked clusters in Polish in several areas of external evidence: written corpora, casual speech as well as in first language acquisition. Following the assumptions underlying markedness, it is expected that the most marked clusters in Polish will be infrequent in the corpora, morphonotactic in nature, as well as frequently reduced in casual speech of adult native speakers and in early first language acquisition. The degree of cluster markedness is evaluated by means of the Net Auditory Distance principle (NAD), which specifies well-formedness conditions for clusters of different sizes and word-positions on the basis of 3 criteria, namely, manner and place of articulation as well as the distinction between an obstruent and a sonorant in a sequence. The results show that the most dispreferred clusters are fairly frequent in written corpora, morphologically-driven only to a certain extent, and more frequently simplified in first language acquisition. In spontaneous speech, however, other linguistic variables, such as frequency, dispreferred CV transitions or the pragmatics of lexical items, may override the preferability criterion.

Keywords: (mor)phonotactics, markedness, frequency, corpora

Introduction: (Mor)phonotactics

Phonotactics is a subbranch of phonology studying permissible sound sequences in language. *Morphonotactics* refers to the interaction between phonotactics and morphotactics (Dressler & Dziubalska-Kołaczyk, 2006) and allows to specify consonant clusters which emerge as a result of the intervention of morphology. Polish, as a strongly inflecting language with rich morphology, is well-known for its rich phonotactic and morphonotactic inventory (Dziubalska-Kołaczyk, 2014; Zydorowicz, Orzechowska, Jankowski, Dziubalska-Kołaczyk, Wierzchoń, & Pietrala, 2016). Polish allows for as many as 4 consonants word-initially, e.g. /pstr-/ in *pstrąg* 'trout', 6 consonants word-medially, e.g. /-ntşzvj-/ *wewnątrzzwiązkowy* 'union-internal' and 5 consonants word-finally, e.g. /-mpstf/ in *przestępstw* , 'crime'-gen.pl. Polish clusters may be phonologically or morphologically motivated. The former are referred to as *lexical* or *phonotactic* clusters, e.g., /pt-/ in *ptak* 'bird', whereas the latter are labeled as *morphonotactic*, meaning that they occur as a result of concatenative or nonconcatentive morphological operations, e.g. /sk-/ in *s+kończyć* 'to end' or /ln-/ vs *lnu* 'linen'-gen.sg. (from *len* 'linen'-nom.sg). Numerous Polish cluster types do not comply with any available scales predicting the goodness of a cluster, i.e. *Sonority Sequencing Generalisation* (Selkirk, 1984), *the Optimal Sonority Distance Principle* (Dziubalska-Kołaczyk, 2002), or *the Net Auditory Distance principle* (Dziubalska-Kołaczyk, 2014; Zydorowicz *et al*., 2016). Plateau or reversed-sonority clusters are attested, e.g. /ss-/ in *ssak* 'mammal' or /wk-/ in *łkać* 'to weep,' and they can be fairly frequent, e.g., word-initial /kt-/ in *kto* 'who' (logarithmic frequency of the word being 5.1 according to Mandera, Keuleers, Wodnicka, & Brysbaert, 2014) or *który* 'which' (logarithmic frequency of the word being 4.9 (Mandera *et al*., 2014)).

The aim of this contribution is the analysis of the most dispreferred clusters with respect to four variables: frequency in written corpora, morphological status of the clusters, and speakers' performance in casual speech and first language acquisition.

The Net Auditory Distance model of phonotactics

The model of phonotactics operating with NAD (cf. Dziubalska-Kołaczyk, 2009, 2014) is part of the Beats-and-Binding Phonology (Dziubalska-Kołaczyk, 2002, 2009). Beats-and-Binding (henceforth B&B) Phonology is a syllable-less theory of phonology, embedded in Natural Phonology. ‘The syllable’ in B&BP is epiphenomenal or emergent due to principled phonotactic forces, referred to as preferences. The phonotactic preferences are responsible for different degrees of intersegmental cohesion which, in turn, determines the behaviour of segments and creates the impression of syllable structure. The rationale behind this model of phonotactics is to counteract the preference for CV which overwhelmingly forces clusters to wither either via reduction or epenthesis/prothesis or at least substitution (CCV→CV, CCV→CVCV, CCV→VCCV). Clusters of consonants persist thanks to auditory contrast and its proper distribution across the word. The principle of perceptual contrast is grounded in the semiotic figure-and-ground principle, on one hand, and in phonetics, on the other.

The Net Auditory Distance (NAD) Principle defines cluster preferability in relation to the position in the word (initial, medial and final). A cluster is preferred if it satisfies a pattern of distances specified by the universal phonotactic preference relevant for its position in the word. For example, for the word-initial cluster C1C2V-, the NAD preference reads: NAD (C1,C2) ≥ NAD (C2,V).

It is believed that auditory (perceptual) distance can be expressed by respective combinations of articulatory features which eventually bring about the auditory effect. An indefinite number of articulatory features as well as detailed acoustic cues would have to be investigated in terms of the degree of their contribution to the overall auditory effect obtained in a cluster. As a reliable starting point, manner (MOA) and place of articulation (POA) are selected, as well as the distinction between an obstruent and a sonorant in a sequence (S/O parameter; the S/O distinction replaced the Lx (voiced/voiceless) distinction used in our previous research). Thus, NAD = |MOA| + |POA| + S/O. Table 1 illustrates the way in which number values are being assigned to particular POAs and MOAs. The values for S/O are 1 or 0.

Table 1. Distances in MOA and POA: Polish

OBSTRUENT			SONORANT					VOWEL		
STOP	FRICATIVE		NASAL	LIQUID		GLIDE				
	AFFRICATE			lateral	rhotic					
5.0	4.5	4.0	3.0	2.5	2.0	1.5	1.0	0		
p b			m			w	w̃	1.0	bilabial	LABIAL
		f v						1.5	labio-dental	
t d	ʦ ʣ	s z	n	l				2.0	(post-)dental	CORONAL
	tʂ dʐ	ʂ ʐ			r			2.3	alveolar	
	ʨ ʥ	ɕ ʑ	ɲ					2.6	alveolo-palatal	
						j	j̃	3.0	palatal	DORSAL
k g		x	ŋ			w	w̃	3.5	velar	
								4.0		RADICAL
								5.0		GLOTTAL

For example, NAD preference for an initial Polish cluster *prV* will be calculated as illustrated below:

CC initial: NAD (C1,C2) ≥ NAD (C2,V)

NAD CC = |MOA1 - MOA2| + |POA1 - POA2| + S/O

NAD CV = |MOA1 - MOA2| + S/O

prV: NAD CC = |5-2|+|1-2.3|+1 = 3+1.3+1 = 5.3

NAD CV = |2-0|+0 = 2+0=2

5.3 > 2; difference 3.3

So, the preference NAD (C1,C2) ≥ NAD (C2,V) is observed since 5.3 > 2. The NAD condition for word-final clusters is as follows: NAD (V, C1) ≤ NAD (C1,C2). NAD allows for formulating two types of predictions. On the one hand, it distinguishes between two sets of clusters, preferred and dispreferred ones, which constitutes a binary qualification into unmarked = preferred = natural vs. marked = dispreferred. On the other hand, it allows to build a hierarchy of preference of clusters ranging from the most preferred to the most dispreferred (least preferred), which constitutes a scalar qualification of relative markedness from the most natural = unmarked = preferred to the heavily marked. This clearly is a strong attribute of this measure, since it reflects the human ability to both categorize and build continua.

The empirical study

Resources

The resources explored in the study include an essential dictionary of Polish (Bartnicka-Dąbkowska & Sinielnikoff, 1999), a written corpus of newspaper texts from *Rzeczpospolita* (used to append the word frequency information to the lexical entries in the dictionary), a spoken corpus (Madelska, 2005) and language acquisition data from the CHILDES database (MacWhinney, 2000; Rose & MacWhinney, 2014; Weist, Wysocka, Witkowska-Stadnik, Buczowska & Konieczna, 1984). The subjects (n=3, 2 male, 1 female) were Bartosz (aged 1;7-1;11), Marta (1;07-1;08) and Kubuś (2;1-2;5).

Defining highly dispreferred clusters

Clusters may be placed on a continuum between the most dispreferred and the most preferred ones. The dispreferred clusters are defined as those which fail to meet the NAD criteria formulated above. They may be dispreferred to various degrees. In the present contribution, we narrow down the analysis to the *most* dispreferred clusters. In order to delineate the area of the most marked sequences, we took the following steps. First we analysed the values of all the clusters occurring in the four corpora. It transpired that the scale for the word-initial context ranged from -5.2 to 5.0. We subdivided the scale into 4 areas: highly dispreferred clusters (henceforth HD), moderately dispreferred clusters (henceforth MD), moderately preferred clusters (henceforth MP), highly preferred clusters (henceforth HP), with a borderline between preferred and dispreferred edges (NAD value = 0). Along the preferred continuum, moderately preferred clusters take the values from 0 (0 included) to 2.5, whereas highly preferred clusters take the values from 2.6 to 5.0. Along the dispreferred continuum, moderately dispreferred clusters are characterized by the values from 0 (0 excluded) to -2.6 whereas highly dispreferred clusters take the values from -2.7 to -5.2. In the word-final context, Polish clusters may take values from -5 to 4.3. Along the preferred continuum, moderately preferred clusters take the values from 0 (0 included) to 2.1 whereas highly preferred clusters range from 2.2 to 4.3. Among dispreferred clusters, moderately dispreferred clusters take the values take from 0 (0 excluded) to -2.5 whereas strongly dispreferred clusters are characterized by NAD values from -2.6 to -5.0. The exhaustive list of HD clusters in the dictionary includes: /ʧʨ ɕʨ db pt ss st ʃʧ sʦ zd sʃ ʃt zʒ ʣb fp fs ft fʦ gd kt sf tk vd vz zv sɕ zʥ fʃ fʧ vʒ gʥ xʨ ʣv sp zb ɕf fɕ fʨ vʥ vʑ ʃk xʃ dʒ ʃp ʧf tʃ dv sk sx tf zg ɕp ʥv ʨf kɕ fx ps xf gʒ kʃ bʒ pʃ/ word-initially and /ɕʨ kt ʃʧ ps st ʃx pʃ ʃt ft ks ɕɲ/ word-finally.

Hypotheses

Four hypotheses are tested. The testing ground consists of word-initial and word-final double clusters.
H 1: Highly dispreferred clusters are predicted to be infrequent in corpora (dictionary and corpus).
H 2: Highly dispreferred clusters are predicted to be morphologically complex to a large extent.
H 3: Highly dispreferred consonant clusters will undergo reduction in spontaneous production.
H 4: Highly dispreferred consonant clusters will undergo reduction in first language acquisition.

Results

Hypothesis 1: Frequency. Table 2 shows general results for word-initial and -final cluster types and word types in the dictionary. In terms of descriptive statistics, the most dispreferred word-initial clusters constitute 42% of all cluster types (61 out of 145) in the dictionary. In terms of the number of lexical entries, that means 45% (955 out of 2137). The distribution of frequencies among cluster types is presented in Figure 1.

The results show that the most dispreferred clusters constitute nearly a half of the dictionary. However, a more detailed analysis reveals that as many as 44 cluster types (which constitutes 72% of HD clusters) have a low or very low dictionary frequency (1-9 representatives). On the other hand, there are only 3 very frequent cluster types: /sp-/, /st-/ and /pʂ-/. /sp-/, /st-/ are generally problematic to account for in many frameworks; additionally in Polish they may be triggered by morphology. Cluster /pʂ-/ may occur in prefixes {prze-}, {przed-}, {przy-} which generate a plethora of derivatives, e.g. *prze+nosić* 'to relocate', *przed+stawić* 'to introduce' or *przy+pomnieć* 'to remind', which accounts for its strong representation in the dictionary.

Table 2. Dictionary frequency

	word-initial context		**word-final context**	
	cluster types	**words**	**cluster types**	**words**
dictionary total	145	2137	54	568
HD	61	955	11	176
	42%	45%	20%	31%

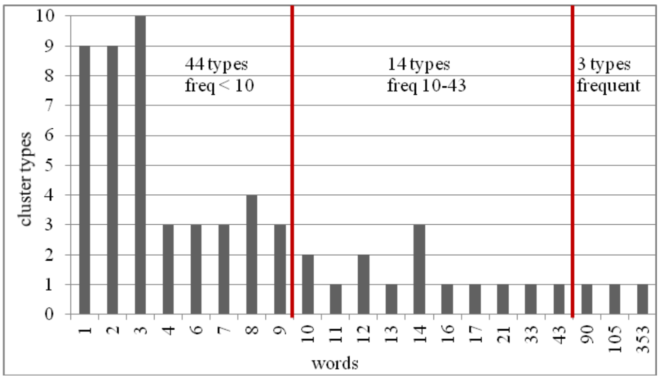


Figure 1. Frequency distribution in the dictionary: Word-initial context

Word-finally, HD clusters constitute 20% (11 out of 54) of all cluster types and 31% words (176 out of 568). The results are presented in Figure 2.

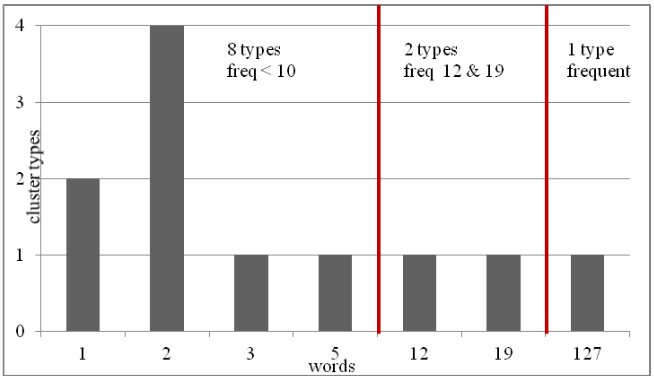


Figure 2. Frequency distribution in the dictionary: Word-final context

Table 3. Corpus frequency

	word-initial context	**word-final context**
corpus total	2 167 149	592 129
HD	1 089 294	219 016
	50%	37%

Similar observations, although on a smaller scale, hold true for consonant sequences in the word-final context. There is only one frequent cluster type, namely /-ɕtɕ/, which may be morphonotactic in nature as in *iś+ć* 'to go' or phonotactic as in *kość* 'bone'. Additionally, it occurs in a very productive suffix {-ość} which forms abstract nouns, e.g. *nowość* 'novelty' or *radość* 'joy'. **Thus the general observation to be made is that among HD clusters there are numerous cluster types with very low frequency and very few cluster types with high frequency.** In the next step, we compare the patterning of HD clusters in the dictionary with the corpus data. The overall frequency as well as the frequency of HD clusters are summed in Table 3.

What needs to be explained at this point is that the corpus frequencies were appended to the lexical entries found in the dictionary (thus if a word is absent from the dictionary, its corpus frequency is not considered either). The data shows that despite their marked character, HD clusters are in common use. Word-initially /pʂ-/ is the top frequent cluster (432 972 tokens, which constitutes 40% (!) of all HD clusters). As mentioned above, cluster /pʂ-/ owes its high frequency to the function of vocabulary items in which it occurs, namely the prefixes {przed-}, {prze-} and {przy-}, but also two prepositions *przed* ‘in front of’ and *przy* ‘at’, which boosts the frequency of the cluster itself. The runner-ups include /sp-, kt-, st-, gd-). Particularly interesting are /kt-/ and /gd-/ sequences as they have a very low dictionary frequency and occur in a handful of words such as *gdy* ‘when’, *gdyby* ‘if’, *gdyż* ‘because’ and *kto* ‘who, *który* ‘which’ (+ derivatives). **This means that both frequent and infrequent cluster types in the dictionary (from two extreme ends of the scale) may boost token frequency in the corpus. Frequency of use clearly depends on the grammar and / or morphological productivity.**

<u>*Hypothesis 2: Morphological structure*</u>. The analysis of the morphological structure of lexical entries in the dictionary revealed that word-initial dispreferred clusters are to a large extent morphologically simple (82% lexical vs 18% morphonotactic word-initially, and 74% vs 26% word-finally). The data, presenting numbers of words, are visualised in Figure 3.

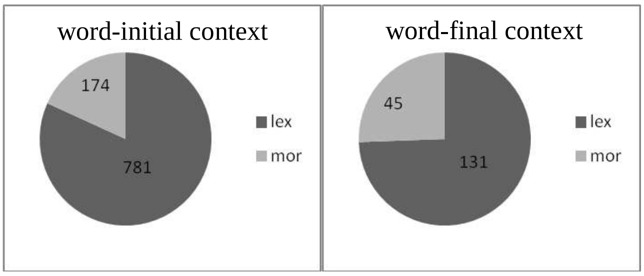


Figure 3. The morphological structure of HD clusters

This goes against the original predictions, therefore, for comparison, we analysed the clusters at the other end of the preferability spectrum, i.e. **strongly preferred ones, which proved to be almost entirely lexical in nature** (with 2% of morphonotactic clusters word-initially and 0% word-finally). The CHI square analysis was run to test the relation between two variables: cluster preferability (at the preferability margins, HD vs HP) and morphological complexity. The two variables turned out to be dependent, the p-value being < .00001. Thus, although there is a relation between preferability and morphological complexity, **the pool of dispreferred clusters is still dominated by lexical clusters**.

<u>*Hypothesis 3: Spontaneous speech*</u>. In order to examine the behaviour of HD clusters in spontaneous speech, we compared their cluster modification rates against the modification rates of all the remaining clusters (MD, MP, HP). The label cluster *modification* rather than *reduction* is used to comprise such processes as consonant deletion, segment substitution and epenthesis. Table 4 presents the data for the word-initial context.

Table 4. Cluster modification rates: Word-initial context

	produced	**reduced**	**total**
HD	3629 (94%)	245 (6%)	3874
remaining	4958 (90%)	522 (10%)	5480
total	8587	767	9354

The CHI square analysis revealed that there is a relation between cluster preferability (HD vs remaining clusters) and cluster modification rates; however, the observation goes *against* the expectations in that it is **strongly dispreferred initial clusters that are simplified less frequently than the remaining groups of clusters** (the p-value is < .00001.). Therefore we conclude that Hypothesis 3 does not hold for clusters in the word-initial context.

This unexpected result deserves a more in-depth analysis. The question to be answered is why less marked or unmarked CC would undergo simplification more frequently. The reasons go beyond the criterion of markedness. Firstly, some preferred clusters occur in words which function as discourse markers, and as such they have a low functional load. Examples include /vw-/ in *właściwie* ‘actually’ and *właśnie* ‘just’, /zr-/ *zresztą* ‘besides’. The second group of clusters particularly susceptible to reduction are those with unfavourable consonant + vowel transitions, e.g. sequences of the type /Cwu/ as is /gw-/ in *głupi* ‘stupid’ or /sw-/ in *słuchać* ‘to listen’. Thus despite being preferred in terms of the NAD distance, they are still vulnerable. Lastly, clusters in high-frequency words are affected, e.g. /vj-/ in *wiem* ‘I know’ or *więc* ‘so’. These 3 criteria may override the unmarked status of a cluster.

Table 5 presents the results of cluster behavior in the word-final context.

Table 5. Cluster modification rates: Word-final context

	produced	reduced	total
HD	263 (26%)	757 (74%)	1020
remaining	454 (77%)	136 (23%)	590
total	717	893	1610

The comparison of HD clusters against the remaining clusters confirmed hypothesis 3: **the most marked final clusters undergo reduction significantly more frequently than the remaining clusters** (the p-value is < .00001.)

<u>*Hypothesis 4: FLA*</u> . In order to examine the behaviour of highly dispreferred clusters in first language acquisition, we compared their cluster modification rates against the modification rates of all the remaining clusters (MD, MP, HP). In this data set, the term cluster *modification* comprises all phenomena affecting clusters, i.e. consonant deletion, vowel epenthesis, vowel prothesis, cluster loss due to word truncation, extension or co-occurrence of the aforementioned processes. Tables 6 and 7 present the data for the word-initial and word-final context respectively (cumulative for all children).

Table 6. Cluster modification rates: Word-initial context

	produced	modified	total
HD	61 (18%)	274 (82%)	335
remaining	158 (39.7%)	240 (60.3%)	398
total	219	514	733

The CHI square analyses, run to test the relation between cluster preferability (HD vs the remaining clusters) and cluster modification rates, revealed that **children indeed perform significantly worse on the most marked clusters** (the p-value being < .00001). Thus we conclude that the assumption of hypothesis 4 has been corroborated.

Table 7. Cluster modification rates: Word-final context

	produced	modified	total
HD	10 (5%)	183 (95%)	193
remaining	13 (26.5%)	36 (73.5%)	49
total	23	219	242

Conclusion

The aim of this contribution was to zoom in on the dynamics of the most marked clusters in Polish. The study reveals that the most marked clusters are by no means rare – they constitute a significant portion of the data both in the dictionary and in the corpus. However, a more precise inspection shows that in the dictionary, there are very few marked cluster types with numerous representatives and numerous marked cluster types with (very) few representatives. HD clusters are also extremely frequent in language use: in the written corpus they constitute 50% of all consonant sequences. However, it must be born in mind that *one* cluster type is responsible for 40% of all HD clusters. As observed above, both frequent and infrequent cluster types in the dictionary (from two extreme ends of the scale) may boost token frequency in the corpus. We conclude that frequency of use clearly depends on the grammar (e.g., a cluster is present in a preposition or other grammatical word) and / or morphological productivity (i.e. a cluster is present in a prefix) of a given cluster rather than on its phonological features (phonotactic preferability). The analysis of the morphological structure of HD clusters revealed that only a subset of the data is generated by morphology. Still, they differ from the highly preferred sequences, which are almost entirely lexical. As regards the behaviour of clusters in casual speech, HD clusters have predictably higher modification rates word-finally and, surprisingly, lower modification rates word-initially in comparison to the remaining clusters. The relative stability of initial marked clusters may be partly due to the word-initial salience principle, partly to the fact that the reducable unmarked clusters are often included in the discourse markers of low pragmatic significance, as well as the fact that the unmarked clusters also include those with auditorily less salient vowel+semivowel sequences (/wu/ and /ji/) which are also susceptible to reduction (this problem would be absent if we excluded clusters with semivowels). Acquisition data, in turn, fully corroborate hypothesis 4: children's performance on the most dispreferred clusters is significantly worse in comparison to the remaining clusters. The preliminary results show that there is a need to further explore the dynamics of phonotactic patterns by including such variables as pragmatic and grammatical functions of words, unfavourable consonant / vowel transitions or token frequency.

References

Bartnicka, B. & Sinielnikoff, R. (1999). *Słownik podstawowy języka polskiego dla cudzoziemców*. Kielce: Wydawnictwo Takt.

Dressler, W. U. & Dziubalska-Kołaczyk, K. (2006). Proposing morphonotactics. *Rivista di Linguistica, 18(2)*, 249-266.

Dziubalska-Kołaczyk, K. (2002). *Beats-and-Binding Phonology*. Frankfurt am Main: Pe-ter Lang.

Dziubalska-Kołaczyk, K. (2009). NP extension: B&B phonotactics, *PSiCL 45 (1)*, 55-71.

Dziubalska-Kołaczyk, K., Pietrala, D. & Aperliński, G. (2014). The NAD Phonotactic Calculator – an online tool to calculate cluster preference in English, Polish and other languages. http://wa.amu.edu.pl/nadcalc/ (date of access: 20 Jul 2017)

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk.* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Madelska, L. (2005). *Słownik wariantywności fonetycznej współczesnej polszczyzny* [Dictionary of phonetic variability in contemporary Polish]. Kraków: Collegium Columbinum.

Mandera, P., Keuleers E., Wodnicka Z., & Brysbaert M. (2014). Subtlex-pl: subtitle-based word frequency estimates for Polish. *Behavior Research Methods 47(2)*, 471-483.

Rose, Y., & MacWhinney, B. (2014). The PhonBank Project: Data and software-assistend methods for the study of phonology and phonological development. In J. Durand, U. Gut & G. Kristoffersen (Eds.), *The Oxford handbook of corpus phonology* (pp. 380-401). Oxford, UK: Oxford University Press.

Selkirk, E. O. (1984). On the major class features and syllable theory. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure*, (pp. 107–136). Cambridge, MA: MIT Press.

Weist, R., Wysocka, H., Witkowska-Stadnik, K., Buczowska, E., & Konieczna, E. (1984). The defective tense hypothesis: On the emergence of tense and aspect in child Polish. *Journal of Child Language*, 11, 347–374.

Zydorowicz, P., Orzechowska P., Jankowski, M., Dziubalska-Kołaczyk, K., Wierzchoń, P. & Pietrala, D. (2016). *Phonotactics and morphonotactics of Polish and English: Description, tools and applications*. Poznań: Wydawnictwo Naukowe UAM.

To cite this article: Zydorowicz, P., & Dziubalska-Kołaczyk, K. (2017). The dynamics of marked consonant clusters in Polish. In E. Babatsouli (ed.), *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017* (pp. 318-324). ISBN: 978-618-82351-1-3. URL: http://ismbs.eu/publications-2017