

**ISEL**  
International Society of Experimental Linguistics

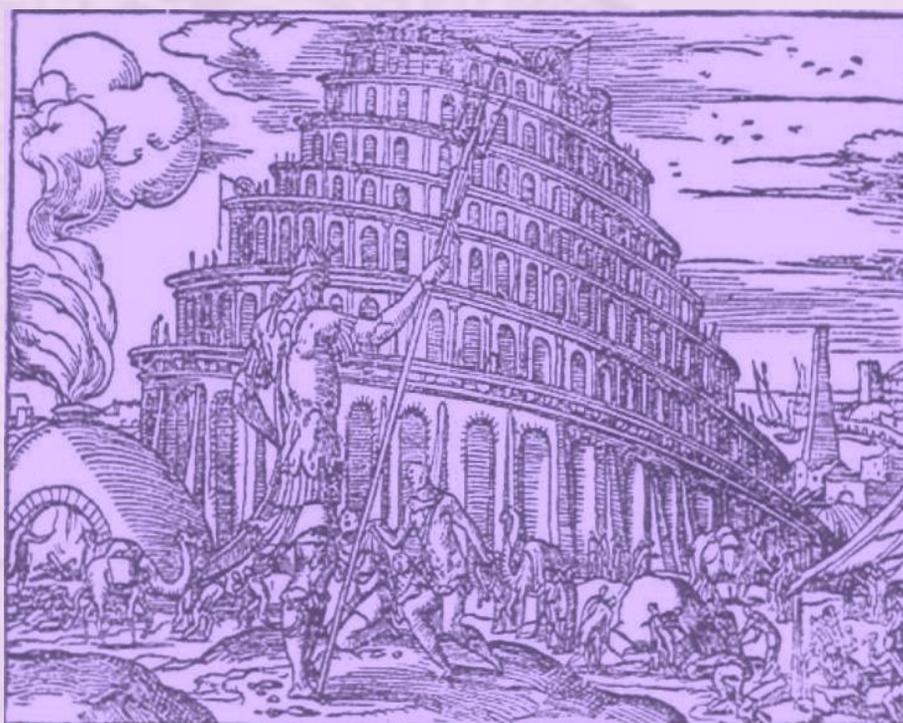
**Proceedings of the 6<sup>th</sup> ISEL Conference on**

# **Experimental Linguistics**

**26 - 27 June 2015**  
**Athens, Greece**



Edited by Antonis Botinis



University of Athens



International  
Speech & Communication  
Association



# **ISEL**

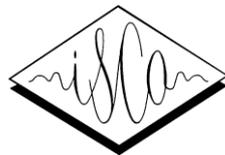
International Society of Experimental Linguistics

Proceedings of the 6<sup>th</sup> ISEL conference

## **ExLing 2015**

26-27 June 2015, Athens, Greece

Edited by Antonis Botinis



International Speech Communication  
Association



University of Athens

Proceedings of the 6<sup>th</sup> ISEL conference  
ExLing 2015

ISEL Editions  
ExLing 2015  
Copyright © 2012 Antonis Botinis

## **Foreword**

---

This volume includes the proceedings of the 6<sup>th</sup> International Conference of Experimental Linguistics, ExLing 2015, held in Athens, Greece, 26-27 June 2015. The first conference was organised in Athens, in 2006, under the auspices of ISCA and the University of Athens and is regularly repeated thereafter.

In accordance with the spirit of this ExLing 2015 conference, we were once again gathered in Athens to continue our discussion on the directions of linguistic research and the use of experimental methodologies in order to gain theoretical and interdisciplinary knowledge.

We are happy to see that our initial attempt has gained ground and is becoming an established forum of a new generation of linguists. As in our previous conferences, our colleagues are coming from a variety of different parts of the world and we wish them a rewarding exchange of scientific achievements and expertise. This is indeed the core of the ExLing events, which promote new ideas and methodologies in an international context.

We would like to thank all participants for their contributions as well as ISCA and the University of Athens. We also thank our colleagues from the International Scientific Committee and our students from the University of Athens for their assistance.

On behalf of the organisation committee

Antonis Botinis

## Contents

<i>Expressive language in 4 to 8 years old children with Down syndrome and typical development: Evidence from the Greek language .....</i>	<i>1</i>
Georgia Andreou, Dimitra Katsarou	
<i>Obstructive sleep apnea syndrome and morphosyntactic development in children .....</i>	<i>5</i>
Georgia Andreou, Matina Tasioudi	
<i>The acoustics of Cypriot Greek fricatives.....</i>	<i>9</i>
Andrie Aristodemou, Angelandria Savva, Charalambos Themistocleous	
<i>Between the distortion and the gradience: Deaf perception of Brazilian Sign Language Hand Shapes.....</i>	<i>13</i>
Felipe Venâncio Barbosa, Janice Gonçalves Temoteo, Roberta Silvestre, Antonieli Cantarelli	
<i>The acquisition of spatial prepositions in Russian .....</i>	<i>17</i>
Maria Grabovskaya	
<i>Testing natural language use: insights from naturalistic experimental paradigms ..</i>	<i>21</i>
Katerina Danae Kandylaki	
<i>L3 acquisition of English present perfect .....</i>	<i>25</i>
Sviatlana Karpava	
<i>Production and perception asymmetries in the Canadian vowel shift.....</i>	<i>29</i>
Thomas Kettig, Bodo Winter	
<i>Question intonation of Greek-American simultaneous bilinguals.....</i>	<i>33</i>
Mary Kontaxi, Anthi Chaida	
<i>Effects of perception and production trainings on the production of English vowels by French native learners.....</i>	<i>37</i>
Jennifer Krzonowski, Emmanuel Ferragne, François Pellegrino	
<i>Computational linguistics &amp; EFL reading comprehension: The KPG text classification profile .....</i>	<i>41</i>
Trisevgeni Liantou	
<i>The production of Spanish vowels by early and late Spanish-English bilinguals .....</i>	<i>45</i>
Natalia Mazzaro, Laura Colantoni, Alejandro Cuza	
<i>Eye movements reflect acoustic cue informativity and statistical noise .....</i>	<i>50</i>
Jessie S. Nixon, Jacolien van Rij, Peggy Mok, Harald Baayen, Yiya Chen	
<i>Cross-category phonological effects on ERP amplitude demonstrate context-specific processing during reading aloud.....</i>	<i>54</i>
Jessie S. Nixon, Jacolien van Rij, Xiaoqing Li, Yiya Chen	

## Contents

---

<i>Realization of Serbian accents by Serbian and Russian speakers</i> .....	58
Ekaterina Panova	
<i>Effects of normal aging on obstruents' acoustic characteristics</i> .....	62
Eirini Papadiamantopoulou, Elina Nirgianaki	
<i>Effect of proficiency on subject-verb agreement processing in French learners of English: An ERP study</i> .....	66
Maud Pélissier, Jennifer Krzonowski, Emmanuel Ferragne	
<i>Perception and production of Italian L2 sounds</i> .....	70
Luciano Romito, Manuela Frontera	
<i>Audiovisual processing of pharyngealization and length in Emirati Arabic</i> .....	74
Mark Scott, Ali Idrissi	
<i>Some effects of Vowel Space Area (VSA) reduction in speech intelligibility</i> .....	78
Patrizia Sorianello	
<i>The intonation of Albanian polar questions and statements</i> .....	82
Charalambos Themistocleous, Daniela Müller	
<i>The influence of tempo and speaking style on timing patterns in Polish</i> .....	86
Agnieszka Wagner, Katarzyna Klessa	

# **Expressive language in 4 to 8 years old children with Down syndrome and typical development: Evidence from the Greek language**

Georgia Andreou, Dimitra Katsarou

Department of Special Education, University of Thessaly, Greece

## **Abstract**

In this paper we report the results of an experiment we carried out in order to investigate expressive language in Down Syndrome (DS). Research has shown that during pre-school and school years, children with DS use expressive skills processes and sound patterns that are similar to those used by typically developing children. However, they eliminate these processes at a slower rate. In view of the above we conducted an experiment so as to investigate expressive language skills in forty five children with DS and forty five children with typical development in order to compare and contrast their language development. Our results showed that children with DS present more expressive language impairments in the Greek language than their typically developing counterparts.

Key words: Down Syndrome, expressive language, Greek

## **Introduction**

Down Syndrome is caused by trisomy of all parts of chromosome 21 and is the most common genetic cause of significant intellectual disability, occurring at approximately 1 in 700 to 1 in 800 live births (Centers for Disease Control and Prevention 2006). Delayed and disordered speech and language development are among the hallmark features of Down Syndrome (Abbeduto and Hagerman 1997). These delays and disorders are usually apparent early in development, but delays in communication and slow rate of language development usually do not cause parental concern, since Down Syndrome can be diagnosed before or at birth (Roberts et al. 2007). Further motor impairments that generally affect the development of speech, including low muscle tone in the tongue, lips and cheeks, result in less firm and precise production of speech sounds (Kumin et al. 1996). As a result, children with Down Syndrome have difficulty in speech production and intelligibility stemming, at least in part, from difficulty with muscle timing and coordination (Iverson et al. 2003). More specifically, previous research in the expressive domain of language between children with Down Syndrome and children with typical development has proven that poor phonological skills could also account for expressive language impairment in Down Syndrome. Numerous studies associate a deficit in expressive skills with DS (Jarrold et al. 2000; Seung and Chapman, 2000). Some research

point out that the problem appears to be independent of hearing impairment or speechmotor deficits (Jarrold et al. 2002; Seung and Chapman, 2000). Nonetheless although some evidence points to lack of articulatory rehearsal as its possible cause (Hulme and Mackenzie, 1992). Some other researchers report that individuals with Down Syndrome might find non-word repetition a particularly difficult task. Tests with words and non-word tasks involve auditory presentation of the non-word stimuli, and spoken repetition responses. Performance is therefore likely to be affected by both hearing ability and speech production skills (Edwards and Lahey, 1998).

## Experimental Procedure

We used a standardized test of expressive language (Tzouriadou et al. 2008) which examines articulation and oral vocabulary. Articulation was examined in two different ways. In the first task, consisting of 13 items, we asked the children to complete a sentence by showing them a picture, such as /ta pedia grafun pano sto.../ (children write on the...) and the children had to answer correctly /thranio/ (desk). In the second task, consisting of 16 items, we asked them to complete the sentence correctly, for example / I Maria ine ginika. O Θomas ine.../ (Mary is a woman. Tom is a ...) and the children had to answer correctly /andras / (man).

Oral vocabulary was examined, also, in two different ways. In the first task consisting of 14 items, we asked the children to find a word that starts with a specific combination of vowels and consonants and means something. For example I want you to tell me a person in the family that is a man and starts with /ba/ and the child had to answer correctly /babas/ (dad). In the second task, consisting of 15 items, we asked them to tell us the definition of a word, for example /ti ine skilos/ (what is a dog?) and the child had to answer /ena zoo pu gavgizi/ (an animal that barks).

## Subjects

45 children with Down Syndrome, mild mental retardation and Greek origin and 45 children with typical development were examined in the two tasks. Their age varied from 4 years to 8 years old and had no other language impairments or health problems that could affect the results of this research.

## Results

For the data analysis we used SPSS version 20.0. A two way analysis of variance (2-way ANOVA) was conducted to determine the effects of age and experimental group on factor *expressive language*. The results indicated a significant main effect for the group ( $F_{7,74} = 3.89$ ,  $p < .05$ ,  $\eta^2 = 0.26$ ), a

significant main effect for age group ( $F_{7,74}=4.29$ ,  $p<.05$   $\eta^2= 0.28$ ) and a significant interaction between the study group and age group. The age group differences for Down syndrome were noted between children aged 48-53 months and 90-95 months ( $MD=1.50$ ,  $p<.05$ ). However in the control group the corresponding differences were noted between children aged 48-53 months and 60-65 months ( $MD=2.11$ ,  $p<.05$ ).

## Discussion

Our findings confirm the continuous problems children with Down Syndrome seem to experience in the expressive domain of language. Difficulties in the expressive language seem to combine with difficulties in articulation, proving that motorspeech deficits can be a serious problem in language development. Children with typical development scored higher in all categories than children with Down Syndrome, an outcome that was expected and is in line with previous research (Yoder and Warren 2004). Expressive language abilities in children with typical development seem to develop in normal age, whereas children with Down Syndrome seem to experience an unusual rate of language development. Children with Down Syndrome aged 48-53 months old seem to develop their expressive language abilities in higher rate than those aged 60-65 months old, proving that expressive language abilities decline as those children enter their school age. This may be explained by the requirements of the test and by the difficulties this specific population experience throughout everyday life at school.

Moreover, the findings of this research show the importance of early language intervention as well as the need for special education to focus on language skills.

However, these results must be treated with caution since it was only a small-scale investigation and children with DS who consisted our sample were not followed longitudinally. Therefore, more longitudinal research studies on Down Syndrome expressive language are needed with large samples which will be followed from childhood right through adolescence and adulthood.

## References

- Abbeduto, L. and Hagerman, R. J. 1998. Language and communication in Fragile X syndrome. *Mental Retardation and Developmental Disabilities Research Reviews* 3, 313-322.
- Centers for Disease Control and Prevention 2006. Improved national prevalence estimates for 18 selected major birth defects-United States, 1999-2001. *Morbidity and Mortality Weekly Report* 54, 1301-1305.

- Edwards, J. and Lahey, M. 1998. Non word repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Linguistics* 19, 279-309.
- Jarrold, C. Baddeley, A. D. and Hewes, A. K. 2000. Verbal short term memory deficits in Down Syndrome: A consequence of problems in rehearsal. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 41, 233-244.
- Jarrold, C. Baddeley, A. D. and Phillips C. E. 2002. Verbal short memory of Down Syndrome: A problem of memory, audition or speech? *Journal of Speech, Language and Hearing Research* 45, 531-544.
- Hulme, C. and Mackenzie, S. 1992. Working memory and severe learning difficulties. Hillsdale, Lawrence Earlbaum Associates Inc.
- Iverson, M., Longobardi, E. and Caselli, C. 2003. The relationship between gestures and words in children with Down Syndrome and typically developing in the early stages of communicative development. *International Journal of Language and Communication Disorders* 38, 179-197.
- Kumin, L., Goodman, M. and Councill, C. 1996. Comprehension speech and language intervention for children with Down Syndrome. *Down Syndrome Quarterly* 1, 1-8.
- Roberts, J.E., Price, J.R. and Malkin, C. 2007. Language and communication development in Down Syndrome. *Mental Retardation and Developmental Disabilities Research Reviews* 13, 26-35.
- Seung, H. K. and Chapman, R.S. 2000. Digit span in individuals with Down syndrome and typically developing children: Temporal aspects. *Journal of Speech, Language, and Hearing Research* 43, 609-620.
- Tzouriadou, M., Singolitou, E., Anagnostopoulou, E. and Vakola, I. 2008. Psychometric criterion of Language Adequacy (L-a-T-o). Thessaloniki, Aristotle University of Thessaloniki and Ministry of Education.
- Yoder, J. and Warren, F. 2004. Early predictors of language in children with and without Down Syndrome. *American Journal of Mental Retardation* 109, 285-300.

# **Obstructive sleep apnea syndrome and morphosyntactic development in children**

Georgia Andreou, Matina Tasioudi

Department of Special Education, University of Thessaly, Greece

## **Abstract**

Obstructive sleep apnea syndrome (OSAS) is a common sleep disorder that ranges up to 3% in childhood epidemiological studies. Language acquisition seems to be affected by OSAS' occurrence but until now research has examined the general language ability without focusing on specific language areas. Therefore, we focused on the morphosyntactic ability of 25 children with OSAS (aged 4.1-6.11) and 25 typically developing children (TDC) of the same age. Performance in morphosyntax was tested by using two subtests (morphosyntactic comprehension and morphosyntactic production) of a standardized language test. Results showed that children with OSAS had significantly lower performance in morphosyntactic production compared to TDC. However, we found no statistically significant difference in morphosyntactic comprehension between the two groups.

Key words: OSAS, children, morphosyntactic development

## **Introduction**

Obstructive Sleep Apnea Syndrome (OSAS) is a common sleep-related breathing disorder in which the upper airway becomes blocked repeatedly during sleep, resulting in increased respiratory effort and snoring, recurrent hypoxemia, and frequent arousals from sleep. Moreover, the prevalence of OSAS in children ranges up to 3% in different epidemiological studies and the incidence peak was found in pre-school children (Gozal 2008). Among the major contributing factors to childhood OSAS, adenotonsillar hypertrophy emerges as the most important condition, while obesity seems to be another basic cause (Alonso-Álvarez et al. 2014). During the first decade of life human brain and neurological functions develop rapidly acquiring the basic cognitive and mental abilities. The sleep disturbances mentioned above lead to increased risk for learning, neurocognitive and behavioral disorders, poor academic performance and school failure (Andreou & Agapitou 2007, among others). Preliminary data suggest that some of these cognitive deficits may be reversible following treatment of sleep apnea in children (Montgomery-Downs et al. 2005); however, factors such as age at treatment time, duration of sleep disordered breathing, pre-morbid intellectual level, socioeconomic status, or the effectiveness of treatment may adversely affect long-term outcome.

The cause of cognitive deficits in OSAS patients seems to be complicated. Some researchers have found a significant correlation between cognitive impairment and daytime sleepiness related to sleep defragmentation, while others attribute cognitive decline to nocturnal hypoxemia. Moreover, it is suggested that the frontal lobes of the brain and the prefrontal cortex are affected by sleep defragmentation and hypoxemia. Another model proposed is the microvascular theory suggesting that a vascular compromise might exist in the small vessels of the brain. The most recent theory suggests that sleep defragmentation and hypoxemia effects are synergistic and they interact with vulnerable brain regions such as hippocampus, prefrontal cortex, subcortical gray and white matter (for a more detailed review see Andreou et al. 2014).

In conclusion, several studies have found that language acquisition is affected by the occurrence of OSAS, but until now research has mostly examined the general language ability without focusing on specific language areas. Therefore, our goal is to investigate a language area that is quite challenging for Greek children with developmental language disorders, namely morphosyntax (Stavrakaki 2005).

## **Experimental Methodology**

In the present study, the participants were 25 children with OSAS, according to polysomnography examination performed at the University Hospital of Larissa, Greece, (Apnea/Hypopnea Index/hour (*AHI*) = 5.25, *SD* = 3.26) aged 4.1 to 6.11 years old (Mean Chronological Age (*MCA*) = 5.6, *SD* = 1.02) and 25 typically developing children (TDC) of the same chronological age (*MCA* = 5.7, *SD* = 1.16/ *AHI* = 0). Performance in morphosyntax was tested using 2 subtests of the standardized psychometric language test L-a-T-o I, created for the Greek language (Tzouriadou et al. 2008). The first subtest examined the morphosyntactic comprehension and consisted of 13 experimental sentences. Three pictures of the same main concept were presented for every experimental sentence and the participant had to choose the one that matched best the sentence heard.

e.g. *Pola skilia kathontai kato*  
Many dogs are sitting down

The second subtest examined the morphosyntactic production and consisted of 13 experimental sentences which were presented orally along with pictures that illustrated the content of each sentence. The examiner pronounced each sentence and stopped at the point where the child needed to add the word or words missing. The missing words were nouns and verbs in singular or plural form, nouns in genitive case, verbs in different tenses (past, present and future) or passive forms.

e.g. Afto to fortigo kouvalai pola koutia. Afto kouvalai... (ena kouti)  
 This truck carries a lot of boxes. This one carries...(one box)  
 We performed independent sample t-tests to compare the performance of the two groups of participants on the tests administered.

**Results**

According to our results, as presented in Table 1, there was a statistically significant difference in the scores for morphosyntactic production between the two groups. No statistically significant difference was found for morphosyntactic comprehension.

Table 1. Mean scores of OSAS and TDC groups in the morphosyntactic tests

Group of participants	Morphosyntactic Comprehension				Morphosyntactic Production			
	<i>M</i>	<i>SD</i>	<i>P</i>	<i>t</i>	<i>M</i>	<i>SD</i>	<i>p</i>	<i>t</i>
OSAS	4.12	1.77	0.024	-2.36	9.88	2.60	0.001	-3.61
TDC	6.45	1.86			11.36	1.21		

**Discussion**

As seen above children with OSAS had significantly lower performance in morphosyntactic production compared to the TDC group. On the other hand, although the OSAS children’s performance was lower than that of the TDC in morphosyntactic comprehension, we found no statistically significant differences between the two groups. Our results are consistent with the literature findings indicating that morphosyntactic production is more challenging than morphosyntactic comprehension in language disorders (Laws & Bishop 2003). Moreover, we found, as expected, that OSAS affects language acquisition in early childhood (Gozal et al. 2008), and we managed to extend the findings of the literature by focusing on a language area that seems to be problematic in children with OSAS, as it has also been found in other developmental language disorders for the greek language (Stavrakaki 2005). We suggest that our findings could result from sleep defragmentation and hypoxemia which affect the function of hippocampus, prefrontal cortex, frontal lobes, and subcortical white matter, since these brain regions are highly associated with language functions and memory (Nagy et al. 2004, Vigneau et al. 2006, among others). Therefore we emphasize that as long as childhood OSAS is concerned medical treatment should be combined with early language intervention in order to avoid the risk of poor language development and the consequent school failure.

## References

- Alonso-Álvarez M.L., Cordero-Guevara J.A., Terán-Santos J., Gonzalez-Martinez M., Jurado-Luque M.J., Corral-Peñafiel J., Duran-Cantolla J., Kheirandish-Gozal L. and Gozal D. 2014. Obstructive sleep apnea in obese community-dwelling children: the NANOS study. *Sleep* 37(5), 943-949.
- Andreou. G. and Agapitou, P. 2007. Reduced language abilities in adolescents who snore. *Archives of Clinical Neuropsychology* 22 (2), 225–9.
- Andreou G., Vlachos F. and Makanikas K. 2014. Effects of Chronic Obstructive Pulmonary Disease and Obstructive Sleep Apnea on Cognitive Functions: Evidence for a Common Nature. *Sleep Disorders*, 1-18.
- Gozal D. 2008. Obstructive Sleep Apnea in Children: Implications for the Developing Central Nervous System. *Semin. Pediatr. Neurol.* 15(2), 100-106.
- Laws G. and Bishop D. V. M. 2003. A comparison of language in adolescents with Down syndrome and children with specific language impairment. *Journal of Speech, Language and Hearing Research* 46, 1324–1339.
- Montgomery-Downs H., Crabtree V.M., Gozal D. 2005. Cognition, sleep and respiration in at-risk children treated for obstructive sleep apnea. *European Respiratory Journal* 25(2), 336-342.
- Nagy Z., Westerberg H. and Klingberg T. 2004. Maturation of white matter is associated with the development of cognitive functions during childhood. *Journal of Cognitive Neuroscience* 16(7), 1227-1233.
- Stavrakaki, S. 2005. Greek neurolinguistics: The State-of-The-Art. *Journal of Greek Linguistics* 5, 181-228.
- Tzouriadou, M., Singolitou, E., Anagnostopoulou, E. and Vakola, I. 2008. Psychometric criterion of Language Adequacy (L-a-T-o). Thessaloniki, Aristotle University of Thessaloniki and Ministry of Education.
- Vigneau M., Beaucousin V., Hevre P.Y., Duffau H., Crivello F., Houde O., Mazoyer B. and Tzourio-Mazoyer N. 2006. Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage* 30(4), 1414-1432.

# The acoustics of Cypriot Greek fricatives

Andrie Aristodemou, Angelandria Savva, Charalambos Themistocleous  
Department of English Studies, University of Cyprus, Cyprus

## Abstract

This study examines the effects of *stress*, *place of articulation* (POA) and *vowel* context on Cypriot Greek voiceless singleton non-sibilant fricative duration, normalised intensity, centre of gravity, standard deviation, skewness, and kurtosis. The findings show significant effects of *POA* on all measurements. Moreover, there were significant effects of *stress* and *vowel* context on duration, centre of gravity, skewness, and kurtosis. To conclude, this study, provides the first instrumental/acoustic data on CG fricatives and advances our knowledge on the effects of *stress* and *vowel* context on the fricatives' acoustic structure.

Key words: fricatives, spectral moments, Cypriot Greek.

## Introduction

Cypriot Greek (CG) contains singletons and geminates voiceless and voiced fricatives, which are articulated in the labiodental /f v/, dental /θ ð/, alveolar /s z/, post-alveolar /ʃ ʒ/, palatal /ç j/, and velar /x ɣ/ place of articulation. Previous research on CG fricatives provides mostly impressionistic evidence (e.g. Newton 1972). Eftychiou (2008) provides additional evidence about the centre of gravity of /s/. This study examines the effects of *stress*, *place of articulation* (POA), and following *vowel* on CG / f θ ç x / duration, normalised intensity, centre of gravity, skewness, and kurtosis.

## Methodology

Twenty-five CG female speakers, born and raised in Nicosia, at their early twenties produced the experimental materials. The later consisted of a set of nonsense words, each containing one of the four fricative ([ f θ ç x ]) in both stressed and unstressed word initial position. In Greek, [ ç ] precedes front vowels /i e/. Specifically, the nonsense words had a CVCV structure: \ 'CVsa (e.g. / 'fasa 'xasa 'çisa/) or CV'sa (e.g., fa'sa, xa'sa, çi'sa, etc./). The carrier phrase was: "/ ipes <keyword> 'pale /" (You told <keyword> again). Each subject produced 64 utterances (i.e. 4 segments × 2 stress placement × 2 vowels × 4 repetitions); a total sum of 1600 productions. Filler words were added in the carrier sentences to provide variation within the experimental material and to minimise speaker's attention on the experimental words. To calculate the acoustic parameters the Discrete Fourier Transformations (DFTs) were averaged using Shadle (2012)'s time-averaging procedure. Within time-averaging, a number of DFTs were taken from across the

duration of the fricative and averaged for each token. The analysed duration of the fricative is always equivalent to the centre 80% of the total duration, cutting off the transitions. Then the first four spectral moments that correspond to the mean (sometimes called the centroid or centre of gravity), standard deviation, skewness, and kurtosis were calculated from the fricative spectra. Linear mixed effects analysis of the relationship between stress, vowel, and segment as fixed effects on the measurements were conducted using R (R Core Team, 2012) and lme4 (Bates, Maechler & Bolker, 2012). As random effects, we had intercepts for speakers and items, as well as by-subject and by-item random slopes for the effect of segment.

## Results

The results about fricatives' duration are reported first and then the results about fricatives' spectral moments: normalised intensity, centre of gravity, skewness, and kurtosis.

Table 1. Mean duration in ms (and SD in parentheses) of fricatives preceding stressed and unstressed /a/ and /i/.

Sounds	Unstressed		Stressed	
	a	i	a	i
f	103 (25)	116 (25)	80 (21)	90 (17)
θ	105 (36)	109 (35)	83 (24)	90 (25)
ç	-	117 (27)	-	95 (20)
x	117 (29)	-	95 (21)	-

*Duration.* Table 1 reports the results from the fricative duration. Stressed fricatives are longer than unstressed fricatives. Unstressed fricatives preceding the vowel /a/ are longer than unstressed fricatives preceding the vowel /i/. Stressed and unstressed /x/ are longer than stressed and unstressed /f/ and /θ/ correspondingly. POA significantly affected duration ( $\chi^2(3)=10.24$ ,  $p < .05$ ); stress also had significant effects on the duration estimate ( $\chi^2(1)=128.64$ ,  $p < .0001$ ) and vowel ( $\chi^2(1)=13.23$ ,  $p < .0001$ ).

Table 2. Mean intensity in dB (and SD in parentheses) of fricatives preceding stressed and unstressed /a/ and /i/.

Sounds	Unstressed		Stressed	
	a	i	a	i
f	31 (8)	34 (9)	34 (9)	34 (9)
θ	33 (10)	33 (10)	35 (9)	33 (10)
ç	-	33 (6)	-	33 (6)
x	37 (4)	-	38 (4)	-

*Normalised Intensity.* Table 2 reports the fricatives' normalised intensity. There was a significant effect of *POA* on normalised intensity ( $\chi^2(3)=10.24, p < .001$ ). *Stress* and *Vowel* had no significant effects.

*Centre of Gravity.* Table 3 reports the fricatives' Center of Gravity. There were significant effects of *POA* ( $\chi^2(3)= 98.40, p < .0001$ ), *Stress* ( $\chi^2(1)= 9.36, p < .05$ ), and *Vowel* ( $\chi^2(1)= 24.94, p < .0001$ ) on the Centre of Gravity.

Table 3. Mean Centre of Gravity (and SD in parentheses) of fricatives preceding stressed and unstressed /a/ and /i/.

Sounds	Unstressed		Stressed	
	a	i	a	i
f	6116 (3877)	7624 (2948)	7231(3722)	8687 (3088)
θ	5798 (3666)	7290 (3527)	6596 (3434)	7901 (3600)
ç	-	6702 (1755)	-	7205 (1905)
x	2607 (720)	-	2841 (861)	-

*Standard Deviation.* Table 4 reports the results mean values and their corresponding SDs for the standard deviation measurements. There was a significant effect of *POA* on standard deviation ( $\chi^2(3)=122.97, p < .0001$ ). *Stress* and *Vowel* had no significant effects.

Table 4. Mean Standard Deviation (and SD in parentheses) of fricatives preceding stressed and unstressed /a/ and /i/.

Sounds	Unstressed		Stressed	
	a	i	a	i
f	4439 (1470)	4710 (1008)	4625 (1413)	4521 (995)
θ	4146 (1310)	4230 (1424)	4688 (1070)	4278 (1279)
ç	-	3618 (697)	-	3591 (536)
x	2487 (990)	-	2818 (927)	-

*Skewness.* There were significant effects of *POA* ( $\chi^2(3)= 81.64, p < .0001$ ), *Stress* ( $\chi^2(1)= 6.37, p < .05$ ), and *Vowel* ( $\chi^2(1)= 17.68, p < .0001$ ) on Skewness.

Table 5. Skewness (and SD in parentheses) of fricatives preceding stressed and unstressed /a/ and /i/.

Sounds	Unstressed		Stressed	
	a	i	a	i
f	1.83 (3.48)	0.29 (1.44)	0.73 (2.3)	-0.09 (1.1)
θ	1.52 (2.48)	0.69 (2.34)	0.92 (1.8)	0.57 (2.4)
ç	-	0.77 (0.98)	-	0.70 (1.34)
x	3.87 (2.13)	-	3.44 (2.22)	-

*Kurtosis.* There were significant effects of *POA* ( $\chi^2(3)= 32.57, p < .0001$ ), and *Stress* ( $\chi^2(1)= 4.06, p < .05$ ) on Kurtosis. *Vowel* also had a minor effect ( $\chi^2(1)= 3.66, p = 056$ ).

Table 6. Kurtosis (and SD in parentheses) of fricatives preceding stressed and unstressed /a/ and /i/.

Sounds	Unstressed		Stressed	
	a	i	a	i
f	16.94 (54)	1.02 (5)	5.33 (16)	0.41 (2)
θ	8.44 (16)	7.32 (23)	3.32 (8)	6.12 (21)
ç	-	1.38 (5)	-	1.48 (8)
x	28.58 (39)	-	20.44 (31)	-

## Discussion

This study investigated the effects of *POA*, *stress*, and *vowel* context on duration, intensity, centre of gravity, standard deviation, skewness, and kurtosis of CG fricatives. The findings show significant effects of *POA* on all measurements. Moreover, there were significant effects of *stress* and *vowel* context on duration, centre of gravity, skewness, and kurtosis. Specifically, stressed fricatives are longer than unstressed fricatives (see also Nirgianaki, 2014). Also stressed and unstressed /x/ is longer than stressed and unstressed /f/ and /θ/ correspondingly. Unstressed fricatives preceding the *vowel* /a/ are longer than unstressed fricatives preceding the *vowel* /i/. Furthermore, *stress* and *vowel* had significant effects on the centre of gravity and skewness but not on standard deviation and normalized intensity. *Stress* has significant effects on kurtosis as well. To conclude, the study provides the first to our knowledge evidence for CG fricatives.

## References

- Armosti, S. (2009). The phonetics of plosive and affricate gemination in Cypriot Greek. (PhD Thesis), University of Cambridge, Cambridge.
- Eftychiou, E. (2009). Lenition Processes in Cypriot Greek. (PhD Thesis), University of Cambridge, Cambridge.
- Newton, Brian (1972). Cypriot Greek. Its phonology and inflections. The Hague, Paris: Mouton.
- Nirgianaki, E. (2014). Acoustic characteristics of Greek fricatives. *Journal of the Acoustical Society of America*, 135(5), 2964-2976.
- Shadle, C.H. (2012). The Acoustics and Aerodynamics of Fricatives. Ch. 20, *The Oxford Handbook of Laboratory Phonology*, eds. A. Cohn, C. Fougerson, M. K. Hoffman, Oxford University Press, pp. 511-526.

# **Between the distortion and the gradience: Deaf perception of Brazilian Sign Language Hand Shapes**

Felipe Venâncio Barbosa<sup>1</sup>, Janice Gonçalves Temoteo<sup>1</sup>, Roberta Silvestre<sup>1</sup>,  
Antonieli Cantarelli<sup>2</sup>

<sup>1</sup>Department of Linguistics, University of São Paulo, Brazil

<sup>2</sup>Institute of Psychology, University of São Paulo, Brazil

## **Abstract**

This study aims observe whether/how Libras Deaf users realize distortions in HS during a Libras' sign accomplishment and whether the hand proximity to the face can influence the perception. A group of deaf signers analyzed a video recorded with signs of Brazilian Sign Language (Libras) performed in two groups: a group of 10 with hands near to the face and far from the face, with three performances, one with the correct sign, one with a slight alteration in finger selection of HS and one distorted. The results suggest that distortion is perceived as a inappropriate expression of a isolated sign. The significant difference between signs with mild distortion near to the face and far from the face can indicate that the linguistic information provided for the face are competitive stimuli and for this reason may inhibit the perception of mild distortions near to the face but may not inhibit the perception of distortions when it is accomplished in a Location without competitive stimuli.

Key words: Brazilian Sign Language, Phonetic-Phonological Analysis, Distortion, Gradience N

## **Introduction**

Hand-Shapes (HS), Location, Movement, Hand Orientation and Non Manual Expressions are described as items that build the internal structure of lexical items in sign languages: the sign (Ferreira-Brito 1995, Quadros e Karnopp 2004, Xavier 2006). In the case of HS, the sort of fingers selection defines specific configurations that compose the sign. Some of these HS are contrastive and delimit minimal pairs in Brazilian Sign Language (Libras). However, the HS accomplishment may present distortions that can be seen as gradient or as errors. The understanding of the accomplishment of phonetic-phonological items in sign language and its perception can bring information about the linguistic processing itself and can give useful information to applied linguistics, mainly in interface areas as speech therapy and neurolinguistics.

This study aims observe whether/how Libras Deaf users realize distortions in HS during a Libras' sign accomplishment and whether the hand proximity to the face can influence the perception.

## Methods

To this end, a video was recorded with an informant performing 19 signs of Libras in two groups: a group of 10 with hands near to the face and a group of 9 signs with hands far from the face. Each sign was accomplished three times, one (correct sign) according to the *New Trilingual Illustrated Encyclopaedic Dictionary of Brazilian Sign Language* (Capovilla, Raphael, & Mauricio, 2009), one (slight distortion) with alterations in finger selection of HS and one (distorted) with great alteration in the finger selection, totalizing 57 presentations. The videos were presented randomly avoiding the sequence of the same signs in the possibilities presented (correct, slight distortion, distorted). Fourteen Deaf informants saw the signs and presented an assessment in a 0 to 10 scale, with marks near to zero to wrong signs and near to ten for correct signs, for each sign.

## Results and Discussion

The responses given in the scales were quantified and the data were analyzed using the ANOVA test to compare the three gradients for each location as well as all of them together (General).

Table 1: Compares Gradients for each Location.

Gradient		Mean	Median	Standard Deviation	VC	N	CI	P-value
Far from the Face	Correct	7.74	10	3.28	42%	126	0.57	<0,001
	Distorted	4.45	4	3.58	80%	126	0.62	
	Slight	4.22	3	3.86	91%	126	0.67	
Near the Face	Correct	8.27	10	3.05	37%	140	0.51	<0,001
	Distorted	4.29	3	3.60	84%	140	0.60	
	Slight	6.31	8	3.98	63%	140	0.66	
General	Correct	8.02	10	3.17	39%	266	0.38	<0,001
	Distorted	4.36	3,5	3.58	82%	266	0.43	
	Slight	5.32	5	4.05	76%	266	0.49	

Table 2: P-values from Table 1.

		Correct	Distorted
Far from the Face	Distorted	<0.001	
	Slight	<0.001	0.651
Near the Face	Distorted	<0.001	
	Slight	<0.001	<0.001
		Distorted	<0.001
		Slight	0.004

Defining the significance level in 5%, results show that the Deaf informants can identify the correct signs and can differentiate it from signs distorted and with slight distortion, with mean scores presenting significant difference ( $p < 0,001$ ), as both group of signs near to the face and signs far from the face. The difference between the general means of signs with mild distortion and signs distorted are statistically significant ( $p = 0,004$ ), however signs with mild distortion accomplished near to the face do not present significant difference when compared with distorted signs ( $p = 0,651$ ). The difference between signs with mild distortion and distorted is significant when accomplished far from the face ( $p < 0,001$ ). The distortion in the fingers selection on HS is identified as inappropriate, although the informant can understand the sign since the HS was distorted and not changed.

We used the same test to compare the two locations in each Gradient. The results shown that there are only significant statistical mean differences among the locations for the Slight Gradient, where we observed a mean of 4.22 for far from the face compared to 6.31 for near the face ( $p$ -value  $< 0001$ ).

Table 3: Compares Locations in each Gradient

Location		Mean	Median	SD	VC	N	CI	P-value
Correct	Far	7.74	10	3.28	42%	126	0.57	0.170
	Near	8.27	10	3.05	37%	140	0.51	
Distorted	Far	4.45	4	3.58	80%	126	0.62	0.705
	Near	4.29	3	3.60	84%	140	0.60	
Slight	Far	4.22	3	3.86	91%	126	0.67	<0.001
	Near	6.31	8	3.98	63%	140	0.66	

The significant difference between signs with mild distortion near to the face and far from the face can indicate that the linguistic information provided for the face are competitive stimuli and for this reason may inhibit the perception of mild distortions near to the face but may not inhibit the perception of distortions when it is accomplished in a Location without competitive stimuli.

## **Conclusions**

The results suggest that distortion is perceived as an inappropriate expression of an isolated sign and it was characterized as a distortion and not as gradient. The significant difference between signs with mild distortion near to the face and far from the face can indicate that the linguistic information provided for the face are competitive stimuli and for this reason may inhibit the perception of mild distortions near to the face but may not inhibit the perception of distortions when it is accomplished in a Location without competitive stimuli.

## **References**

- Capovilla, FC, Raphael, WD, Maurício, ACL. 2009. New Trilingual Illustrated Encyclopaedic Dictionary of Brazilian Sign Language. São Paulo. Edusp.
- Ferreira-Brito, L. 1995. Por uma gramática da Língua de Sinais. Rio de Janeiro: Editora Tempo Brasileiro.
- Quadros, RM, Karnopp, LB. 2004. Língua de Sinais Brasileira – estudos lingüísticos. Porto Alegre: ArtMed.
- Xavier, AN. 2006. Descrição fonético-fonológica dos sinais da língua de sinais brasileira. Dissertação de mestrado. São Paulo, USP.

# The acquisition of spatial prepositions in Russian

Maria Grabovskaya

National Research University Higher School of Economics, Russian Federation

## Abstract

In this work the acquisition of prepositions *v* ‘in’, *na* ‘on’, *u* ‘near’, *pod* ‘under’, *nad* ‘above’ and *za* ‘behind’ was tested. Children between ages 3:0,6 and 7:1,9 were asked to describe everyday situations based on pictures (e.g. *sobaka v budke* ‘the dog in the kennel’). The results show that the prepositions *v* ‘in’ and *pod* ‘under’ are acquired first, followed by the preposition *na* ‘on’. Then the prepositions *nad* ‘above’ and *za* ‘behind’ are acquired.

Key words: Language acquisition; Spatial prepositions.

## Introduction

In Russian both prepositions and cases are used to express the location of the object. In this work only the acquisition of prepositions is examined. Cases start to appear earlier than prepositions (Gvozdev 1961, Ufimceva 1979). After the appearance of cases “stopgaps” and “pseudoprepositions” (vowel sounds or real prepositions) begin to emerge (Berman 1985, Leikin 1989), although this stage is not found in the speech of every child (Leikin 1989). It is also important to mention that before prepositions children start using locative adverbs (Gvozdev 1961, Leikin 1989). The meaning of prepositions is understood in all the situations, not in the particular contexts (Clark 1973, Johnston 1988, Wilcox & Palermo 1982).

In this article we are going to test the hypothesis of Leikin (1998), who studied the ability of Russian children between ages 3;0 and 7;2 to produce and understand locative expressions, and suggested that the acquisition of locative prepositions is not complete until approximately the age of 7. Figure 1 represents the direction of substitutions in naming spatial relations established by Leikin based on experimental tasks that involve manipulation of geometrical objects. We look at prepositional use for more common situations of speech production.

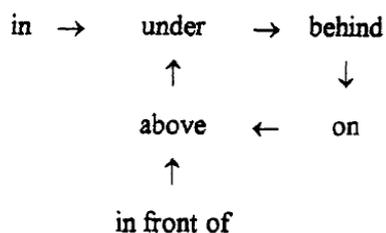


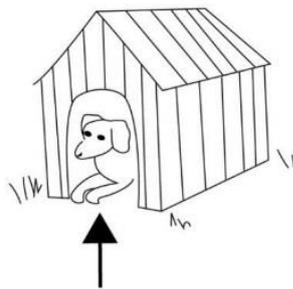
Figure 1. “Directions of substitutions in naming spatial relationships (most frequent types of substitutions)” (Leikin 1998).

## Method

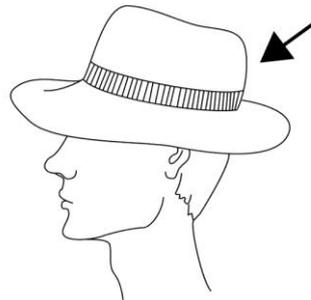
### Subjects

Responses were elicited from 57 subjects between the ages of 3:0,6 and 7:1,9. 29 subjects from all the children were male, 28 subjects – female. For all the children Russian was a native language. 8 children were 3 years old, 9 children were 4 years old, 23 children were 5 years old, 16 children were 6 years old and there was 1 child who was 7 years old.

### Procedure



Picture 1. *Sobaka v budke* ‘the dog in the kennel’.



Picture 2. *Shljapa na golove* ‘the hat on the head’.

As stimuli, we have 24 + 4 pictures from Topological relations picture series (Bowerman & Pederson 1992) where 4 pictures were used for training. All the pictures depict everyday situations. Each picture was presented separately. From the main 24 pictures 2 pictures were supposed to provoke the preposition *v* ‘in’ in the description of these pictures; 2 pictures were supposed to provoke the preposition *pod* ‘under’, 1 picture – the preposition *u* ‘near’, 2 pictures – *nad* ‘above’, 17 pictures – *na* ‘on’. There were more pictures for the last mentioned preposition because they hold different types of meaning. Compare the expected phrases: *kniga na polke* ‘the book on the shelf’, *treshhina na chashke* ‘the crack on the cup’, *jabloko na jablone* ‘the apple on the apple-tree’. We analyzed the results for all the 24 pictures and 1 picture from 4 training pictures (that is supposed to provoke the preposition *za* ‘behind’).

The order of the main 24 pictures was the same for all the children. The task for the children was to describe the picture starting with the object to which the arrow points. For example, for Picture 1 expected answer is *sobaka v budke* ‘the dog in the kennel’, for Picture 2 – *shljapa na golove* ‘the hat on the head’. Sometimes the interviewers used the hints to help a child (e.g. naming of the difficult object). The linguists didn’t help children with the prepositions.

## Results

The process of acquiring prepositions is not finished for all our subjects, which corresponds to the hypothesis of Leikin (1998). The prepositions *v* ‘in’ and *pod* ‘under’ are acquired earlier and used correctly even by children of 3 years old, as predicted by Leikin (1998). The preposition *na* ‘on’ is acquired a bit later (approximately at the age of 4). The preposition *u* ‘near’ appeared to have many synonyms (*rjadom s*, *vozle*, *okolo*) but this doesn’t contradict its early acquisition but points at its low frequency. The prepositions *nad* ‘above’ and *za* ‘behind’ are acquired later. Moreover, *nad* ‘above’ seems to be acquired earlier than *za* ‘behind’. Some stimuli appeared to be hard to describe despite the fact that the expected preposition was relatively basic (example: *jabloko na strele* ‘the apple on the arrow’). This picture is often described in other ways – e.g. *jabloko protknuto streloj* ‘the apple picked by the arrow’. The results of Gvozdev (1961) and Leikin (1989) were also proved: locative adverbs are acquired earlier than prepositions. A child tends to describe the picture with adverbs (for example, *naverhu* ‘up’, *vnizu* ‘down’) when he or she doesn’t have the needed preposition acquired (smth is situated up and smth is situated down). Children also tend to replace some of the nouns that are part of the object on the object itself: e.g. replacement of *jabloko na vetke* ‘the apple on the branch’ on *jabloko na dereve* ‘the apple on the tree’.

## Notes

The children were asked by the author and a group of other linguists who also needed the data but for their own different work about other aspects of speech acquisition.

## Acknowledgements

The author would like to thank Tatyana Nikitina for helpful advices during the work on this project.

## References

- Berman, R. A. (1985). The acquisition of Hebrew. In D. Slobin (Ed.), *The crosslinguistic Study of language acquisition*, 255—371. New York: Erlbaum.
- Bowerman, M., & Pederson, E. 1992. Topological relations picture series. In S. C. Levinson (ed.), *Space stimuli kit 1.2*, 51. Nijmegen: Max Planck Institute for Psycholinguistics.
- Clark, H. 1973. Space, time, semantics and the child. In T. Moore (ed.), *Cognitive development and the acquisition of language*, 27-64. New York: Academic Press.

- Gvozdev, A. D. 1961. Voprosy izucheniya detskoy rechi [Problems of studying children's speech]. Moscow, Izdatelstvo Akademii Pedagogicheskikh Nauk RSFSR.
- Johnston, J. J. 1988. Children's verbal representation of spatial location. In J. Stiles-Davis, M. Kritchevsky, & U. Bellugi (eds.), *Spatial cognition*, 195-205. Hillsdale, NJ: Erlbaum.
- Leikin, M. W. 1989. Puti logopedicekoj raboty po preodoleniu agrammatizmov v ponimanii i upotreblenii prostransivennykh predlogov u bolnykh s aphasiy [Speech therapy on prepositional agrammatism in aphasia]. Unpublished doctoral dissertation, Leningrad, Leningrad State, Pedagogical Institute.
- Leikin, M. 1998. Acquisition of Locative Prepositions in Russian. *Journal of Psycholinguistic Research* 27, 91-108.
- Ufimceva, N. V. 1979. Psikholingvisticheskie kharakteristiki funkcionirovanija kategorii padezha sushestvujushhego v russkom jazyke [Psycholinguistic Issues of Grammar]. In F. Tarasov (ed.), *Psikholingvisticheskie problemy grammatiki*, 5-65. Moscow, Moscow University.
- Wilcox, S. & Palermo, D. 1982. Children's use of lexical and non-lexical information in responding to commands. *Journal of Child Language*, 9, 139-50.

# Testing natural language use: insights from naturalistic experimental paradigms

Katerina Danae Kandylaki

Department of Germanic Linguistics, University of Marburg, Germany

## Abstract

Experimental approaches to language processing should pay special attention to natural language use. Traditional experimental paradigms use controlled, carefully created stimuli, simplified by decontextualisation. Real language use, however, is full of elliptical, ambiguous sentences where meaning inference rests on contextual cues, and requires not only structural and semantic interpretation for memory storage, but is used to inform actions and interactions. Cognitive neuroscience research is slowly starting to include natural language use in experimental paradigms, by using so called naturalistic paradigms. The current paper presents a naturalistic approach to test language processing and compares results from the current approach to traditional approaches. In conclusion, insights from naturalistic paradigms can inform linguistic theory on the basis of human brain processes.

Key words: naturalistic paradigms, online language processing, fMRI

## Two traditions of experimental linguistics

Inspired by experimental sciences and by theoretical and descriptive linguistics, psycho- and neurolinguists often formulate hypotheses that are based on theoretical categorisations of linguistic systems. This, however, might not mirror the way in which our brain categorises language.

The “controlled stimuli” tradition follows strict design principles from psychological science. Researchers create or select carefully controlled stimuli that usually range from single words to maximally two-sentence passages. Then they manipulate certain aspects of the stimuli, according to their hypotheses e.g. the 'b' vs. 'd' distinction in phonology, active vs. passive voice in syntax and test the hypotheses by calculating differential contrasts. In these designs, the participants are usually asked to deliver some judgement (e.g. for the 'b' vs. 'd' distinction, they might be asked to state if the presented word is a pseudoword) or respond to a probe (e.g. as in questions “Was this word in the list you encountered previously?”). Nuisance factors are eliminated as far as possible from the design by the careful control of presentation (randomisation of trials) and by the existence of enough fillers. Traditional statistical approaches (t-tests, chi-squared tests) are applied to analyse the results. In the end, if the experiment was diligently constructed and carried out, the conclusions can be drawn with respect to the

initial hypotheses. It is, however, an open question if such findings can scale up to allow conclusion about real language use.

The necessity of the alternative approach of the “ecological validity” tradition is underlined by the inevitable abstraction of controlled experimental paradigms, which create an experimental reality, far removed from our everyday linguistic reality. This tradition has been boosted increasingly in the last 15 years by new methodological advances in cognitive neuroscience (e.g. Bartels & Zeki, 2004; Hasson, Nir, Fuhrmann, & Malach, 2004) which allow us to present more complex stimuli. In the “naturalistic” paradigms (Willems, 2015), participants are presented with stimuli that capture natural language use as in reading (e.g. Kurby & Zacks, 2012, 2013), listening to stories (e.g. AbdulSabur et al., 2014), or even communicating in a dialogue (e.g. Kuhlen, Allefeld & Haynes, 2012). These approaches are characterised by greater freedom in participants' behavior and loss of strict experimental control. In one example experimental situation, a pair of participants is given a problem and they have to interact in order to solve it. What is measured can be the common lexical elements that the two participants use and whether the number of these elements increases as progress is made towards the solution in order to achieve a more efficient collaboration (alignment in dialogue). At the very end of the naturalistic approaches continuum lie completely data-driven experiments as in Bartels & Zeki (2004). In this study, the results are not modeled by an a priori predictor model, but are derived automatically from the fMRI data by using an algorithm that identifies spatial and temporal activation dynamics during “free-viewing” of a movie.

Although these two traditions might initially seem impossible to combine, recent attempts have been made to combine ecologically valid designs in which controlled stimuli are “hidden” (see for example Haupt, Schlesewsky, Roehm, Friederici, & Bornkessel-Schlesewsky, 2008). In this way, we take advantage of a controlled experimental design and can still draw conclusions that have a higher probability of applying to natural language use.

### **The current approach**

In more detail, the current approach uses a story listening paradigm in which participants listen to 20 stories in a functional magnetic resonance imaging (fMRI) experiment. The stories provide a rich auditory language stream of information which allows for a novel manipulation of several linguistic factors in the same natural context. After each story, its content was tested in two comprehension questions. The participants' task was to listen to the stories carefully and answer the two subsequent questions after each story. For the purposes of the current paper we would like to elaborate mainly on

the stimuli and one part of the analyses, in order to highlight the advantages of a combined naturalistic approach.

The stories include the following linguistic manipulations: 1. phonological rhythm processing, 2. syntactic and semantic cues for reference tracking and 3. discourse processing of false-belief passages. We managed to embed highly controlled experimental conditions in this naturalistic setting by including the manipulations of our factors and filling the rest of the stories context that was as natural as possible. Each of our factors constrained the stimuli at a different level 1. specific compound words for the manipulation of rhythm processing, 2. specific transitive verbs for the reference tracking question, 3. specific discourse for the false belief situations, where one person in the story needed to have a false belief (Frith & Frith, 1999).

When analysing the data we aimed at describing as many processes as possible, in order to achieve a clean baseline. Therefore, similarly to a mixed effects modelling approach in which random factors are added in order to regress out as much noise as possible, we modelled: 1. the speech within the stories which was irrelevant to each manipulation, 2. question reading, 3. the response and 4. inter- and intra-trial jitters.

In the current paper we report results on the rhythmic processing manipulation as an example of testing natural language use. Domahs, Klein, Huber, & Domahs (2013) have conducted a controlled study with similar research questions and found activation for rhythmic processing in the following brain regions: bilateral superior temporal gyrus (STG), right precuneus, left angular gyrus, bilateral inferior frontal gyrus (IFG), bilateral supplementary motor area (SMA), left insula. We found comparable brain regions in a similar manipulation: bilateral STG, right IFG, bilateral SMA and left insula. Moreover, we found additional brain regions which support rhythmic processing in language in context: left premotor cortex (PMC), right post-central gyrus. In the Domahs et al. (2013) study participants were instructed “to decide whether the auditory probe was stressed correctly and as expected given the visual target”. In our study participants only listened to the stories without specifically being instructed to pay attention to the stress pattern of the stimuli. Our results therefore establish the ecological validity of the Domahs et al. (2013) study and add to the previously identified rhythm processing regions the PMC and post-central gyrus.

## **Conclusion**

Naturalistic designs are necessary for the enrichment of experimental linguistic research, if our goal is to understand language performance. The way in which the brain processes and analyses linguistic input might not be captured in our current theories of linguistic knowledge which are created by

reflecting on linguistic output. The processing of language adds useful insights to the system of linguistic knowledge and can inform applications of linguistic interaction.

### **Acknowledgements**

I am grateful to my advisor, Prof. Dr. Ina Bornkessel-Schlesewsky, for her close guidance in this project. This research was financed by the ExInit initiative and by the LOEWE project “Research focus: exploring fundamental linguistic categories” of University of Marburg.

### **References**

- AbdulSabur, N. Y., Xu, Y., Liu, S., Chow, H. M., Baxter, M., Carson, J., & Braun, A. R. 2014. Neural correlates and network connectivity underlying narrative production and comprehension: A combined fMRI and PET study. *Cortex*, 57, 107-127.
- Bartels, A., & Zeki, S. 2004. Functional brain mapping during free viewing of natural scenes. *Human brain mapping*, 21(2), 75-85.
- Domahs, U., Klein, E., Huber, W., & Domahs, F. 2013. Good, bad and ugly word stress—fMRI evidence for foot structure driven processing of prosodic violations. *Brain and language*, 125(3), 272-282.
- Frith, C. D., & Frith, U. 1999. Interacting minds – a biological basis. *Science*, 286(5445), 1692-1695.
- Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. 2004. Intersubject synchronization of cortical activity during natural vision. *science*, 303(5664), 1634-1640.
- Haupt, F. S., Schlesewsky, M., Roehm, D., Friederici, A. D., & Bornkessel-Schlesewsky, I. 2008. The status of subject–object reanalyses in the language comprehension architecture. *Journal of Memory and Language*, 59(1), 54-96.
- Kuhlen, A. K., Allefeld, C., & Haynes, J. D. 2012. Content-specific coordination of listeners' to speakers' EEG during communication. *Frontiers in human neuroscience*, 6.
- Kurby, C. A., & Zacks, J. M. 2012. Starting from scratch and building brick by brick in comprehension. *Memory & cognition*, 40(5), 812-826.
- Kurby, C. A., & Zacks, J. M. 2013. The activation of modality-specific representations during discourse processing. *Brain and language*, 126(3), 338-349.
- Willems, R. M. (Ed.). 2015. *Cognitive neuroscience of natural language use*. Cambridge University Press.

# **L3 acquisition of English present perfect**

Sviatlana Karpava

University of Central Lancashire, Cyprus, Cyprus Acquisition Team

## **Abstract**

This study investigates L3 acquisition of English present perfect by Greek Cypriot speakers. One hundred CG university students took part in the study, the first part of which examined the sensitivity to grammatical norms (a passage correction task, based on Odlin et al., 2006), and the other part was focused on the production of English present perfect (elicitation of natural discourse, essays about personal experience). The results showed that L3 learners used more non-target tense forms (present simple and past simple) than the target present perfect in obligatory contexts, which is due to transfer from L1 CG. The findings are in line with the Typological Primacy Model (Rothman, 2010), as L3 learners transfer from L1 and this transfer is negative.

Key words: present perfect, L3 acquisition, L1 transfer, pragmatic conditions

## **The Study**

According to the Typological Primacy Model (TPM) (Rothman, 2010), proximity in actual or perceived linguistic typology is the most important in adult multilingual syntactic transfer (Rothman and Cabrelli Amaro, 2007; Rothman, 2010). Both L1 and L2 can be a source for transfer in L3/Ln, dependent upon typological and psychotypological similarities between target L3 and L1 or L2. Psychotypology is the speaker's perception of typological proximity between languages (Kellerman, 1983). TPM is in line with the Full Transfer/Full Access Hypothesis. Transfer occurs selectively and it can be both positive and negative.

With regard to current research, semantic and pragmatic conditions compatible with present perfect are different in CG, SMG and English, so it is expected that L3 learners would fail to notice these conditions. They might have a tendency to use past simple instead of existential present perfect due to transfer from L1 CG. They might overlook these semantic and pragmatic conditions related to present perfect, as in their L1 there are no such meanings and conditions and, as a result, would equate the semantics of present perfect with semantics of past simple. The aim of this study is to reveal the cause of non-target/deviant production of L3 English present perfect, the direction and the source of transfer to L3, and the role of semantic/pragmatic contexts of the present perfect, lexical aspect, transitivity of the verb, type of the sentence, type of the adverbial modification as well as the role of age, age of onset and the length of L3 input on the comprehension and production of English present perfect.

## Study

100 CG university students (89 undergraduate, 11 postgraduate) took part in the study (69 males and 31 females, age: 17-36 years, length of exposure to L3 input: 2-20 years, age of onset to L3: 10 to 27 years). The first part of the study examined the sensitivity to grammatical norm (a passage correction task), while the other part of the study was focused on the actual production of English present perfect (elicitation of natural discourse). The proofreading test based on Odlin et al. (2006) had 60 test items (25 errors: present perfect, resultative and existential, replaced with present simple and past simple and 35 distractors: correct and incorrect usage of present simple/continuous, past simple/continuous, future simple).

## Results and Discussion

The analysis of the error correction task showed that only 400 (16%) of all errors were corrected and L3 learners used target present perfect, the other errors (2100/84%) were either not corrected or L3 learners tended to use other non-target tense forms instead of present perfect: past simple (1154/46.16%), past perfect (6/0.24%), past continuous (36/1.44%), present simple (809/32.36%) or present continuous (95/3.80%). L2 learners had more 'no changes' for past simple test items (63.76%) than for present simple test items (48.41%), with both past simple and present simple test items having the same percentage of changes to present perfect (17% and 16.33%). Present simple test items were more changed to past simple (29%) than past simple test items were changed to present simple (17.33%). Consequently, the most preferable tense used instead of present perfect is past simple, see Table 1:

Table 1. (No) attempted corrections of test items.

	Past simple items	Present simple items
No changes	829 (63.76%)	581 (48.41%)
Changed to		
Present perfect	221 (17%)	196 (16.33%)
Past perfect	7 (0.53%)	3 (0.25%)
Present simple	225 (17.33%)	
Present continuous	18 (1.38%)	47 (3.93%)
Past continuous		25 (2.08%)
Past simple		348 (29%)

No significant difference was revealed between target production for existential and perfective present perfect. But L3 learners of English used

more past simple for existential present perfect (50.44%) than for resultative present perfect (38.70%). This can be due to transfer from L1 CG (usage of past simple instead of existential present perfect). They used more present simple for resultative present perfect (39.60%) than for existential present perfect (27.46%). These findings are in line with the Typological Primacy Model (Rothman, 2010), as L3 learners transfer from L1 CG. This transfer is negative, non-facilitative, and CG is psychologically perceived to be typologically closer to English (than SMG to English) due to the post-colonial situation in Cyprus and widespread usage of English on the island. Overall, L3 learners showed better production for distractor items than for test items. This suggests that they have a particular problem with present perfect rather than with other tenses (present simple/continuous, past simple/continuous, future simple). They had a higher percentage for acceptance of the correct distractor items (75.85%) than for the correction of incorrect distractors (52.34%).

A paired samples t-test showed a statistically significant difference between target and non-target present perfect production ( $t(99)=14.992$ ,  $p=.000$ ), target present perfect and non-target past simple production ( $t(99)=8.060$ ,  $p=.000$ ), target and non-target distractor production ( $t(99)=9.338$ ,  $p=.000$ ), and present perfect and past simple production in existential contexts ( $t(99)=8.713$ ,  $p=.000$ ). One-way ANOVA showed that age, length of exposure to L3, and gender are not crucial factors for L3 present perfect production. Pearson correlation analysis showed that target and non-target present perfect production is correlated with target and non-target distractor production (proficiency): Sig 2-tailed .000. Thus, L3 proficiency is the crucial factor for target/non-target L3 present perfect production.

With regard to elicitation of natural discourse: discourse about personal experiences based on essays, overall, it was very difficult to elicit present perfect in natural discourse due to the low rate of present perfect production (151 obligatory present perfect contexts for 100 essays). It was found that L3 learners used more non-target tense forms (64.91%): past simple (45.05%) or present simple (19.86%), than target present perfect (35.09%) in the obligatory present perfect contexts. They used both present and past simple instead of present perfect due to similarities of certain features of present perfect and present simple (current relevance) and present perfect and past simple (anterior) (Bardovi-Harlig, 1997).

It was found that target present perfect was used mainly in resultative contexts, while non-target past simple was used both in resultative and experiential/existential contexts, and non-target present simple was used in resultative, extended-now and recent past. It seems that the semantic context of present perfect influences target and non-target production of present

perfect in L3 English. Cypriot Greek students tend to use past simple instead of present perfect in existential/experiential contexts, which can be explained by L1 transfer, as in CG they use past simple instead of experiential/existential present perfect. Target present perfect was mainly used with achievement verbs, non-target past simple was used both with achievement and state verbs, and non-target present simple was used with achievement, state and activity verbs. The data supports the Inherent Lexical Aspect Hypothesis (Andersen and Shirai, 1996; Bardovi-Harlig, 1999), as L3 learners use mainly achievement and accomplishment verbs with perfective and past tense morphology.

This study is an attempt to shed light on multilingual development — L3 acquisition of English in Cyprus with regard to present perfect. Both comprehension and production of this particular linguistic phenomenon have been examined. It was found that L3 learners transfer from L1 CG rather than from L2 SMG, specifically using past simple instead of existential present perfect. L3 learners ignore semantic and pragmatic conditions compatible with the use of English present perfect; they mostly equate the semantics of the past tense with the semantics of the present perfect. This is in line with the Typological Primacy Model (Rothman, 2010), as this transfer is non-facilitative and there is also a (psycho)-typological proximity between CG and English.

## References

- Andersen, R. and Shirai, Y. 1996. Primacy of aspect in first and second language acquisition: The pidgin/creole connection. In T.K. Bhatia and W. Ritchie (eds.), *Handbook of Second Language Acquisition*, 527-570. London, Academic Press.
- Bardovi-Harlig, K. 1997. Another piece of the puzzle: The emergence of the present perfect. *Language Learning* 47, 375-422.
- Bardovi-Harlig, K. 1999. Exploring the interlanguage of interlanguage pragmatics: A research agenda for acquisitional pragmatics. *Language Learning* 49, 449-65.
- Kellerman, E. 1983. Now you see it, now you don't. In Gass, S. and Selinker, L. (eds.), *Language Transfer in Language Learning*, 112-134. Rowley, MA, Newbury House.
- Odlin, T., Alonso, R. and Alonso-Vazquez, C. 2006. Fossilization in L2 and L3. In Han, Z.H. and Odlin, T. (eds.), *Studies of Fossilization in Second Language Acquisition*, 56-83. Clevedon, England, Multilingual Matters.
- Rothman, J. 2010. On the typological economy of syntactic transfer: Word order and relative clause attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics (IRAL)* 48, 243-71.
- Rothman, J. and Cabrelli, J. 2007. On the initial state of L3 (Ln) acquisition: Selective or absolute transfer? Paper presented at the 5th International Conference on Third Language Acquisition, Stirling, Scotland, September.

# Production and perception asymmetries in the Canadian vowel shift

Thomas Kettig<sup>1</sup>, Bodo Winter<sup>2</sup>

<sup>1</sup>Theoretical and Applied Linguistics, University of Cambridge

<sup>2</sup>Cognitive and Information Sciences, University of California, Merced

## Abstract

The Canadian Shift (CS), a lowering and backing of the KIT, DRESS, and TRAP vowels, has been extensively investigated in the speech of English Canadians. However, its effect on the perceptual categorization of vowels has received little attention. In general, the role of perception in ongoing vowel shifts remains comparatively under-researched. By testing participants both in production and in perception, this study gives a unique view into an ongoing sound change. Participants from Montreal in two age groups were recorded reading a list of 44 sentences containing words with stressed /ɛ/, /æ/, /ɔ/, and /ʌ/ vowels. Participants also categorized 96 synthesized vowel stimuli. While the production data clearly replicated the CS, in perception, shift-leaders did not categorize vowels very differently.

Key words: Canadian Vowel Shift; English; sound change; speech perception

## Introduction

The Canadian Shift (CS) describes a systematic lowering and backing of the KIT, DRESS, and TRAP vowels. However, Boberg (2008: 137) concludes in an overview of the Canadian English vowel system that “the regional profile of the Canadian Shift is far from clear.” This paper follows up on earlier work from Boberg (2005) and addresses the apparent-time trajectory of the CS in a community of English speakers in Montreal. We do not only investigate the ongoing sound in speech production, but also in speech perception (cf. Kendall & Fridland 2012; Thomas 2002), allowing us to see the extent to which speakers who are leading the change in production are also leading the change in perception.

## Methodology

To control for the effect of local ethnolect, this study is limited to the Jewish community of Montreal (pop. 80,000). Our group contained 12 young speakers (born 1984 or later, 5 female, 7 male) and 16 older speakers (born 1961 or earlier, 5 female, 11 male).

Each participant performed two experiments. In Experiment I, participants spoke a list of 44 sentences containing words with stressed /ɛ/, /æ/, /ɔ/, and /ʌ/ vowels. The target words were controlled for voicing, place

of articulation, and manner of articulation of the consonant following the vowel. In Experiment II, we presented human-sounding synthetic vowels to participants, each 250ms in duration. In each trial, four large buttons of equal size, with the labels BAT, BET, BUT, and BOUGHT, appeared on a screen for participants to select as they listened to 96 stimuli, which formed an F1 continuum from 700 Hz to 950 Hz and a F2 continuum from 1200 Hz to 1950 Hz. Stimuli were played in the same, random order for every participant.

## Results

Figure 1 displays the Lobanov-normalized vowel space, with each vowel symbol representing the mean F1/F2 of one speaker, plus confidence ellipses. Age differences are most pronounced for /æ/ and /ɛ/. Both are retracted for younger speakers. The orientation of the ellipses shows that the change of /æ/ is primarily one in F2. On the other hand, the change of /ɛ/ is expressed both in F1 and F2, although it is more pronounced for F2. Linear mixed effects models confirm these observations: There are significant age differences in F2 for /æ/ ( $p < 0.001$ ) and /ɛ/ ( $p = 0.03$ ), but not for /ʌ/ and /ɔ/ ( $p > 0.1$ ). There was also a nearly significant effect in F1 for /ɛ/ ( $p = 0.06$ ).

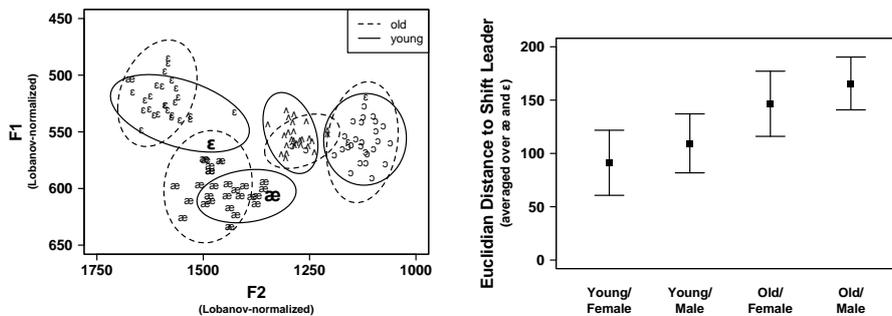


Figure 1. Means for each speaker's vowels (production), plus 95% confidence ellipses (solid line for younger speakers, dotted line for older speakers).

Figure 2. Linear mixed effects model predictions for Euclidian distance to shift leader (error bars indicate 95% confidence interval).

Based on Figure 1, we selected a speaker whose /æ/ mean was most retracted, as well as a speaker whose /ɛ/ mean was most lowered. Since these two speakers have the most extreme values in the direction of the shift, they can be thought of as “shift leaders,” highlighted in boldface in Figure 1. We used the two speakers' mean F1/F2 values of /æ/ and /ɛ/ to calculate Euclidian distances for each vowel. This Euclidian distance can be interpreted as indicating a measure of “shift leadingness” for every vowel. Figure 2 displays this measure for the age/gender groups in our data. A



binary partitioning on each participant's /æ~ε/ F2 categorization curves. As can be seen, this measure correlates with age, and there is a significant effect of age (permutation test,  $p=0.036$ ). Younger speakers accept more retracted vowels (lower F2) as realizations of /æ/. In other words, the point at which they switch to /ε/ occurs significantly later on the F2 continuum.

## Discussion

Production results demonstrate that /æ/ and /ε/ are shifting in apparent time in the vowel spaces of English-speaking Montrealers (Figure 1). An analysis of inter-speaker variation shows a significant degree of ordered heterogeneity, with young women leading the change and older males retaining the most conservative pronunciations, the typical progression for a sound change advancing below the level of consciousness. Though Boberg's (2005) study of Montreal found /ε/ to be retracting and /æ/ to be both lowering and retracting, it seems as though the operation of the CS in Montreal now involves the retraction of /æ/ without any accompanying lowering, while /ε/ is simultaneously backed and somewhat lowered. One implication of this finding could be that over the last decade, /æ/ could have lowered as far in the vowel space as it will go, and is now only retracting (cf. Roeder & Jarmasz 2010).

Although the categorization behaviour between /æ/ and /ε/ is significantly affected by age (but only for F2), there are overall less pronounced differences in perception than in production. In fact, those vowels that do not show age differences in production, /ʌ/ and /ɔ/, show the biggest age differences in perception. This finding suggests that in an ongoing change, speakers' productions may differ, but their perception needs to accommodate the fact that both new and old variants need to be heard: younger speakers need to understand older speakers who do not participate as strongly in the shift. Thus, in the Canadian Vowel Shift, both production and perception are changing, but production changes more than perception.

## References

- Boberg, C. 2005. The Canadian Shift in Montreal. *Language Variation and Change* 17: 133-154.
- Boberg, C. 2008. Regional Phonetic Differentiation in Standard Canadian English. *Journal of English Linguistics* 36: 129.
- Kendall, T. & Fridland, V. 2012. Variation in perception and production of mid front vowels in the US Southern Vowel Shift. *Journal of Phonetics* 40: 289-306.
- Roeder, R. & L.-G. Jarmasz. 2010. The Canadian Shift in Toronto. *The Canadian Journal of Linguistics* 55(3): 387-404.
- Thomas, E. R. 2002. Sociophonetic Applications of Speech Perception Experiments. *American Speech* 77(2): 115-147.

# Question intonation of Greek-American simultaneous bilinguals

Mary Kontaxi<sup>1</sup>, Anthi Chaida<sup>1,2</sup>

<sup>1</sup>Dept of Speech and Language Therapy, AKMI Metropolitan College, Greece

<sup>2</sup>Lab of Phonetics & Computational Linguistics, University of Athens, Greece

## Abstract

This study examines the influence of American English language on the tonal characteristics of Greek polar and wh-questions. Three Greek-American simultaneous bilinguals and three Greek monolinguals participated in a production experiment. The results indicate that Greek polar questions are not radically affected by American English tonal patterns, except for the span and alignment of the boundary tone. The tonal influence of American English is more apparent in wh-questions, which present a different boundary tone. Our findings suggest that simultaneous bilinguals may use tonal features of both languages, which are not however consistent in all question types.

Key words: bilingualism, intonation, questions, Greek, American English

## Introduction

The present study focuses on the influence of American English language in the tonal characteristics of Greek polar (yes/no) and wh-questions produced by Greek-American simultaneous bilinguals. Since intonation is one of the most essential communication constituents, carrying both linguistic and paralinguistic features (Botinis et al. 2001), the investigation of the under-examined issue of L1-L2 tonal interference, especially with regards to Greek language, may reveal important aspects for both communication research and language acquisition.

Bilinguals are considered as individuals who have acquired the knowledge and the use of two languages. However, the degree of bilingual proficiency, the context of language acquisition, the age of acquisition, the circumstances under which the languages are used (i.e. domains of use) and, the social orientation, constitute the 5 most important variables related to bilingualism (Baker 2011). Focus has been given on the cross-linguistic interference in several languages (e.g. German-English, Turkish-German, German-Spanish), in order to detect the degree of the influence of both languages in simultaneous bilinguals' linguistic competence (Döpke 2000; Gut 2000). It is evident that bilinguals' intonation patterns in one language have been affected by the other language to some extent (Queen 2001; Lleó & Rakow 2011).

## Method

Three Greek-American simultaneous bilingual and three monolingual of Standard (Athenian) Greek female speakers, within the age range of 50-65 years, participated in a production experiment. All bilinguals were born in the US and lived there on average of 13,6 years, have been exposed to both languages from their birth until 3 years of age, and have received education (preschool, elementary, middle and high school) in English, while Athens has been their permanent residence for 42,3 years on average.

The experimental material consisted of 2 simple SVO sentences. The sentences were produced in 3 renditions: neutral statement, polar question and wh-question, without focus, and without any lexical or morphosyntactic alteration, apart from the necessary addition of the word /ja'ti/ (why) for the production of the wh-question (Table 1). The final corpus included 72 utterances (3 speakers X 2 sentences X 3 renditions X 2 repetitions).

Table 1: The speech material of target sentences and their context.

Sentence Type	Context	Target Sentence
Statement	ti 'jinete? (What is going on?)	i 'vana ya'zoni ma'di'la (Vana is sewing handkerchiefs)
		i 'mina mi'razi ba'loŋa (Mina is distributing balloons)
Polar question	ti me 'rotises? (What did you ask me?)	i 'vana ya'zoni ma'di'la? (Is Vana sewing handkerchiefs?)
		i 'mina mi'razi ba'loŋa? (Is Mina distributing balloons?)
Wh-question	ti me 'rotises? (What did you ask me?)	ja'ti i 'vana ya'zoni ma'di'la? (Why is Vana sewing handkerchiefs?)
		ja'ti i 'mina mi'razi ba'loŋa? (Why is Mina distributing balloons?)

## Results

The results show that polar questions follow a similar tonal structure in both monolingual and bilingual productions, except for the span and alignment of the boundary tone. In bilingual productions, a rise-fall movement covers the 2 last syllables of the utterance, while the peak of the boundary tone is aligned with the end of the stressed vowel of the penultimate syllable. In monolingual productions a distinct and sharper rise-tonal fall forms the boundary tone, which only covers the ultimate syllable of the utterance (Figure 1). This tonal difference of the 2 last syllables of bilingual and monolingual polar questions is significant ( $F(1,142) = 22.255, p < 0.05$ ).

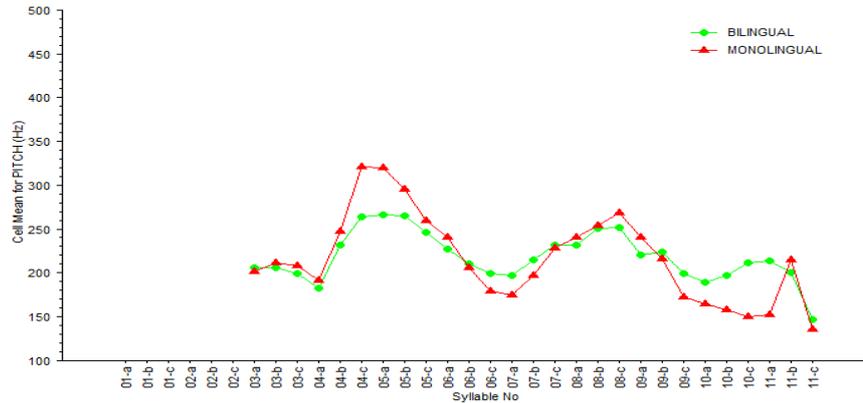


Figure 1: Tonal structures (mean values in Hz) of polar questions of monolingual and bilingual speakers.

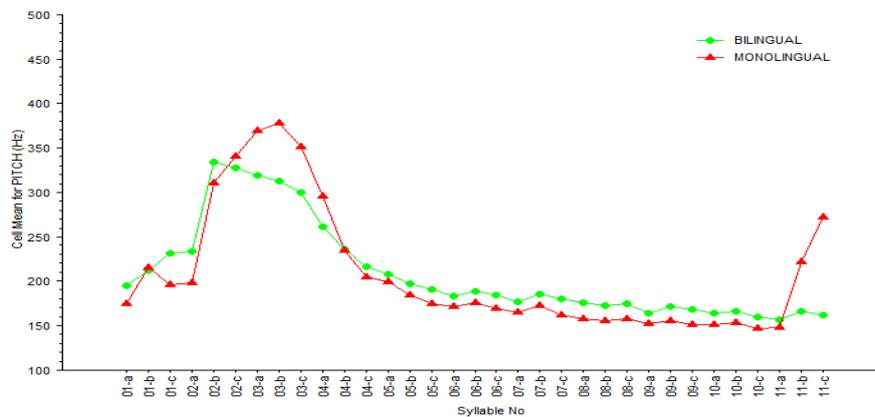


Figure 2: Tonal structures (mean values in Hz) of wh-questions of monolingual and bilingual speakers.

A more noticeable difference is observed in wh-questions. Although in both cases there is a distinct focal accent covering the wh-word and the 2 following syllables, in bilinguals' productions the tonal peak is aligned with the beginning of the stressed syllable (2<sup>nd</sup>) of the wh-word, while in monolinguals' productions the tonal peak is sharper and slightly moved to the right, aligning with the vowel of the 3<sup>rd</sup> syllable. Moreover, a different boundary tone is noted. In monolinguals' productions the boundary tone is a typical tonal rise at the last syllable, while in bilinguals' productions a low plateau beginning from 5<sup>th</sup> syllable covers the whole utterance until the end (Figure 2). This tonal difference of the 2 last syllables of bilingual and monolingual wh-questions is significant ( $F(1,142) = 5.487, p < 0.05$ ).

## Conclusion

In accordance with the results of this study, American-English does not radically affect the intonation patterns of Greek-American bilinguals in Greek polar questions, since bilinguals' productions followed a similar pattern with the monolinguals' ones. However, there is a noticeable American English interference in the case of Greek wh-question productions by bilinguals, since their boundary tone resembles the falling boundary tone pattern of American English wh-questions (Hedberg et al. 2010; Kainada & Lengeris 2013), rather than the typical rising pattern of Greek wh-questions (Botinis 2011; Chaida 2010).

Our findings suggest that simultaneous bilinguals of Greek and American English may use tonal features of both languages, which are not however consistent and global in all sentence types.

## References

- Baker, C., 2011. *Foundations of Bilingual Education and Bilingualism*. 5<sup>th</sup> ed. Bristol: Multilingual Matters.
- Botinis, A., 2011. *The Phonetics of Greek* (in Greek). 2<sup>nd</sup> ed. Athens: ISEL editions.
- Botinis, A., Granström, B. and Möbius, B., 2001. Developments and Pragmatics in Intonation Research. *Speech Communication* 33, 263-296.
- Chaida, A., 2010. Production and Perception of Intonation and Sentence Types in Greek. PhD thesis, University of Athens.
- Dopke, S., 2000. Introduction: On the Status of Cross-Linguistic Structures in Research on Young Children Growing Up With Two Languages Simultaneously. In: Dopke, S., ed. *Cross-Linguistic Structures in Simultaneous Bilinguals*. Amsterdam: John Benjamins, pp.1-11 .
- Gut, U., 2000. Cross-Linguistic Structures in the Acquisition of Intonational Phonology by German-English Bilingual Children. In: Dopke, S., ed. *Cross-Linguistic Structures in Simultaneous Bilinguals*. [online] Amsterdam: John Benjamins, pp.201-226.
- Hedberg, N. and Sosa, J.M., 2002. The Prosody of Questions in Natural Discourse. Proceedings of Speech Prosody 2002. Aix-en-Provence, France, 11-13 April 2002, 375-378.
- Kainada, E. and Lengeris, A., 2013. The acquisition of English intonation by native Greek speakers. Proceedings of the 20th International Symposium on Theoretical and Applied Linguistics. Thessaloniki, Greece.
- Lleo, C. and Rakow, M., 2011. How do German-Spanish bilingual children ask wh-questions in their two languages? *Proceedings of the 17<sup>th</sup> International Congress of Phonetic Sciences*. Hong Kong, China, 17-21 August 2011.
- Queen, R.M., 2001. Bilingual intonation patterns: Evidence of language change from Turkish-German bilingual children. *Language in Society*, vol. 30, 55-80.

# Effects of perception and production trainings on the production of English vowels by French native learners

Jennifer Krzonowski<sup>1</sup>, Emmanuel Ferragne<sup>2</sup>, François Pellegrino<sup>1</sup>

<sup>1</sup>Laboratoire Dynamique du Langage UMR5596 CNRS/Université de Lyon, France

<sup>2</sup>CLILLAC-ARP EA3967, Université Paris Diderot, France

## Abstract

This study examined the effect of two different trainings on the production of English vowels (/i:/, /ɪ/, /æ/, /ʌ/ and /ɑ:/) by French learners. Forty-eight French first-year students, who had learned English in school only, were divided into three groups receiving either 5 sessions of perceptual training (PE-Group), or 5 sessions of production training (PR-Group) or no training (C-Group). They were recorded at pre-test and post-test with a reading task of /bVd/ words, and their performance was evaluated by discriminant analysis based on sex-specific models trained on native speakers productions. The results show improved classification rates for the vowels /ʌ/ and /ɪ/ in the PR-Group and only for /i:/ in the PE-Group. No improvement was observed for the C-Group.

Key words: L2 phonology acquisition, vowel production, training

## Introduction

Phonetic trainings have already shown their efficiency for improving perception and production performance for second-language learners both for consonant (Bradlow, 2008, Iverson, Hazan and Bannister, 2005, Pruitt, Jenkins and Strange, 2006) and vowel categories (Aliaga-Garcia, 2010, Iverson and Evans, 2009, Iverson, Pinet and Evans, 2011, Lambacher, Martens, Kakehi, Marasinghe and Molholt, 2005). Among the studies carried out on the acquisition of second language vowels, very few relate to the acquisition of English vowels by French learners and performance assessment is often restricted to perception. Iverson, Pinet and Evans (2011) tested the effect of a classical High Variability Perceptual Training on the acquisition of English vowels by French speakers. Their results in production showed significant improvement only for /i:/, /ɛI/ and /ɑY/.

Our study investigates the effect of two different types of training (perception and production) on the production of English vowels by French learners. We focused on two regions of the vowel space that are known to cause difficulties for these learners: /i: - ɪ/ and /æ - ʌ - ɑ:/. These vowels are located in the vowel space where there is only one category in French but two (/i: - ɪ/) or three (/æ - ʌ - ɑ:/) in English.

## Procedure

The study followed a pre-test / training (five 1-hour sessions on consecutive days) / post-test paradigm. The test sessions included a production task involving three /bVd/ words for each vowel category.

Forty-eight first-year students of English were divided into 3 groups: one group received a high variability perceptual training (PE-Group)(Iverson, Pinet and Evans, 2011), a second one received a production training (PR-Group). A last control group (C-Group) received no training but listened to audiobooks in English between test sessions.

The training material consisted of recordings by native speakers of British English. Ten CVC minimal pairs for each of the three contrasts involved (/i:/-/ɪ/, /æ/-/ʌ/ and /ʌ/-/ɑ:/) were produced by twelve native speakers (6 women) from the South East of England.

The first three training sessions involved one of the three contrasts each, and the last two sessions combined all contrasts. For the PE-Group, the first three sessions contained 2AFC identification and AX discrimination tasks with feedback, while the last two sessions were composed of 5AFC identification and oddity discrimination tasks with feedback.

The PR-Group performed a word repetition task with feedback in all sessions. The feedback was visually given from a real-time acoustic analysis of the participants' production (vowel formants and duration) compared to those of the native speakers of the same sex. The result of an automatic classification of the vowel based on a sex-specific model trained on native English productions was also displayed.

## Data analysis

Our expectations were that, if the training was effective, automatic vowel classification would improve between the pre-test and the post-test. The first two formant values measured at vowel midpoint were computed, Bark-transformed, and z-scored (independently for F1 and F2) for each speaker (Ferragne and Pellegrino, 2010). The natives' formant values were used to train a linear discriminant classifier with one model for women and one for men. Then the learners' vowels were automatically classified based on the appropriate model.

## Results

The overall correct classification score for the pre-test is 78%. Important discrepancies between vowel categories are found: /i:/: 71%, /ɪ/: 77%, /æ/: 92%, /ʌ/: 55%, /ɑ:/: 94%. Clearly, most misclassifications come from the vowel /ʌ/. Classification rates for all vowel categories taken together improved between pre-test and post-test for the PR-Group (7 percentage

points - pp) and the PE-Group (7 pp), both reaching statistical significance according to a binomial test (respectively,  $p = 0.02$  and  $p < 0.01$ ), whereas it dropped for the C-Group (-6 pp,  $p = 0.03$ ).

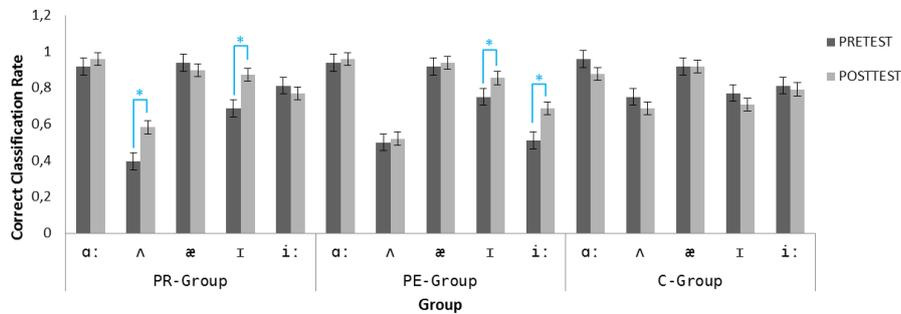


Figure 1. Classification rates for all vowels at pre-test (dark bars) and post-test (light bars) for each group with error-bars.

Now focusing on the vowels (see Figure 1), classification rates for /ʌ/ increased by 19 pp for the PR-Group, which was found to be a significant improvement ( $p = 0.007$ ). For the PE-Group, the increase was only 2 pp (not significant). The C-Group decreased by 6 pp (not significant). The vowel /ɪ/ was also significantly better categorized at post-test for the PR-Group, increasing by 18 pp ( $p = 0.006$ ). The PE-Group's performance also increased for this vowel (10 pp,  $p = 0.012$ ). As for /ʌ/, at pre-test, the vowel /ɪ/ was often misidentified as /ʌ/, and less often so at post-test. The PE-Group showed another significant improvement for the vowel /i:/ (18 pp,  $p = 0.014$ ).

## Conclusion

There is therefore evidence that the production training was effective for the acquisition of /ʌ/ but not the perception training. Both training groups improved on the production of /ɪ/, but only the PE-Group performed better for the entire /i: - ɪ/ contrast. These results will be compared to perceptual scores in a follow-up study. The different patterns observed between the training groups could be correlated with the perception patterns for these two contrasts: it may be that the /i:/-/ɪ/ contrast is more difficult to perceive than /æ/-/ʌ/ and that a perceptual training is more appropriate for the former contrast to help French speakers.

## Acknowledgements

This research was supported by an IUF grant to Emmanuel Ferragne and by the ASLAN Labex.

## References

- Aliaga-Garcia, C. 2010, Measuring perceptual cue weighting after training: A comparison of auditory vs. articulatory training methods, K. Dziubalska-Kołaczyk, M. Wrembel and M. Kul (eds.), *New Sounds 2010*, Poznan, Poland
- Ferragne, E. and Pellegrino, F. 2010, Formant frequencies of vowels in 13 accents of the British Isles, *Journal of the International Phonetic Association*, 40, 1-34
- Iverson, P. and Evans, B. G. 2009, Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers, *Journal of the Acoustical Society of America*, 126, 866-877
- Iverson, P., Hazan, V. and Bannister, K. 2005, Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults, *Journal of the Acoustical Society of America*, 118, 3267-3278
- Iverson, P., Pinet, M. and Evans, B. G. 2011, Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels, *Applied Psycholinguistics*, 33, 145-160
- Lambacher, S. G., Martens, W. L., Kakehi, K., Marasinghe, C. A. and Molholt, G. 2005, The effects of identification training on the identification and production of American English vowels by native speakers of Japanese, *Applied Psycholinguistics*, 26, 227-247
- Pruitt, J. S., Jenkins, J. J. and Strange, W. 2006, Training the perception of Hindi dental and retroflex stops by native speakers of American English and Japanese, *Journal of the Acoustical Society of America*, 119, 1684-96.

# **Computational linguistics & EFL reading comprehension: The KPG text classification profile**

Trisevgeni Liontou  
Greek Ministry of Education

## **Abstract**

Advances in Computational Linguistics and Machine Learning systems have made it possible for EFL teachers, material developers and test designers to go beyond surface text components and adopt more theoretically sound approaches to text readability, focusing on a wider range of deep text features. Taking advantage of recent developments, the present research explored the existence of any statistically significant lexicogrammatical differences between intermediate and advanced reading comprehension exam texts of the Greek State Certificate of English Language Proficiency national exams in order to better define text complexity per level of competence. The main outcome of the research has been the Text Classification Profile that includes a qualitative and quantitative description features pertinent in intermediate and advanced reading comprehension exam texts.

Key words: linguistics, text complexity, reading comprehension

## **Introduction**

The present study builds on earlier findings of research on reading assessment, according to which many text variables such as content, lexis and structure can have an impact on either the reading process or product and need to be taken into account during test design and validation (Oakland & Lane 2004).). In fact, although a lot of research has been conducted in the field of second language acquisition with specific reference to ways of reading and text processing strategies, Alderson (2000) stressed language testers' lack of success "to clearly define what sort of text a learner of a given level of language ability might be expected to be able to read or define text difficulty in terms of what level of language ability a reader must have in order to understand a particular text". Such information would be particularly useful in providing empirical justification for the kinds of reading texts test-takers sitting for various language exams are expected to process, which to date have been arrived at mainly intuitively by various exam systems (Fulcher 1997).

### **Aim of the study**

The aim of the present study was to delineate a range of linguistic features that characterize the reading comprehension texts used at the B2 (Independent User) and C1 (Advance User) level of the Greek national exams in English for the State Certificate of Language Proficiency -known with their Greek acronym as KPG exams- in order to better define text readability per level of competence and create a statistical model for assigning levels to texts in accordance with the purposes of the specific exam battery. In order to explore these issues, the following research question was formed:

1. Is there a specific set of text variables to better predict text difficulty variation between reading texts used at the B2 and C1 KPG English language exams?

### **Methodology**

Over the last ten years, advances in Computational Linguistics and Machine Learning systems have made it possible to go beyond surface text components, focusing on a wider range of *deep* text features that take into account semantic interpretation and the construction of mental models and can, thus, offer a principled means for test providers and test-takers alike to assess this aspect of test construct validity (27). In the present study *Coh-Matrix 2.1*, *Linguistic Inquiry and Word Count 2007* (LIWC), *VocabProfile 3.0*, *Computerized Language Analysis* (CLAN) suite of programs, *Computerized Propositional Idea Density Rater 3.0* (CPIDR), *Gramulator*, *Stylometrics* and *TextAnalyzer* were used to estimate 135 text variables.

### **Text Analysis Findings**

The ultimate purpose of the present research was the creation of a mathematical formula that could make possible the classification of English texts to two levels of language competence, that is, intermediate (B2) or advanced (C1), according to their level of linguistic complexity. In order to avoid contamination of results due to text length variation, the frequency counts of all indices were normalized to a text length of 100 words. IBM SPSS 20.0 statistical package data was used to compute descriptive statistics and perform reliability analyses, Pearson product moment correlations, T-tests, ANOVAs, multiple Linear and Binary Logistic Regressions.

In order to create our statistical model, regression analysis, a procedure commonly employed for predictive purposes, was used. The training set consisted of 63 KPG reading texts, 34 of which had already been used in past B2 level exams and 29 in C1 level exams. These texts had originally

been chosen as appropriate for each level of competence based on the judgement of experienced KPG test designers and a series of piloting sessions. Due to the high number of independent text variables and the two levels of text classification Binary Logistic Regression was carried out using the Forward method and the Wald criterion in IBM SPSS 20.0. Forward selection looks at each explanatory variable individually and selects the single explanatory variable that fits the data the best on its own as the first variable included in the model. Given the first variable, other variables are examined to see if they will add significantly to the overall fit of the model. Among the remaining variables, the one that adds the most is included. This latter step (examining remaining variables in light of those already in the model and adding those that add significantly to the overall fit) is repeated until none of the remaining variables would add significantly or there are no remaining variables (Cook *et al.* 2000). Following this method the program entered in the analysis one variable at a time, depending on whether it met a predefined set of statistical criteria ( $p < 0.05$ ), and also removed those that contributed less to the predictive power of the model. To avoid over-fitting our model, a minimum of 15 cases of data for each predictor variable was considered acceptable, with the final model containing 4 predictor variables. Thus, with a ratio of 15.75, the model followed the necessary statistical restrictions and distanced itself from over-fitting problems (Foster 2001). Based on the model, the resulting regression equation or prediction formula, which may provisionally be called the **L.A.S.T. Text Difficulty Index** is as follows:

$$\begin{aligned} \text{Predicted Text Level} = & 23.097 + (0.202 * \text{LexicalDensity}) \\ & + (0.814 * \text{AcademicWordList}) \\ & + (-97.184 * \text{SyntacticSimilarityAll}) \\ & + (-60.381 * \text{Tokens/Family}) \end{aligned}$$

The specific regression model succeeded in correctly predicting the level of 33 of the 34 B2 texts (97%) and 26 of the 29 C1 texts (90%). The total percentage of correct predictions on the training set was 94%. This result signifies that the combination of the four variables alone managed to correctly classify 59 of the 63 B2 and C1 KPG reading texts used in the analysis. In other words, for the training set, using these specific four variables, the model has managed to correctly predict the level of 94% of the total number of pre-classified passages. The Lexical Density (LD) variable refers to the proportion of content words to the total number of running words in a text estimated through LIWC2007, the Academic Word List variable relates to Coxhead's academic list and was estimated through VocabProfile, the Syntactic Structure Similarity variable for all sentences

across paragraphs was calculated using Coh-Metrix 2.1, whereas the proportion of Tokens per Family was provided by VocabProfile.

## Concluding remarks

The preliminary results of the present study showed that the L.A.S.T. Text Difficulty Index could be used to draw a rough distinction between intermediate and advanced texts based on four linguistic features, i.e. lexical density, syntactic structure similarity, tokens per word family and technical vocabulary. This application might prove useful to test developers and other stakeholders interested in automatic text classification since texts calibrated to specific levels of language competence, can be fed into an electronic bank, from which test task writers of pen-and-paper or e-tests can more consistently choose source texts on the basis of specific text attributes. In addition, the proposed model might be of value in the context of classroom-based assessment, as well as for other exam batteries to better define what sort of text a reader of a prospective level of language ability should be able to process under real exam conditions. Undoubtedly, the new formula could be best viewed as a springboard for assessing text difficulty based on a more sound theory of language. For now, it remains to be explored whether the L.A.S.T. Text Difficulty Index will perform as well with a wider range of reading texts from various disciplines.

## References

- Alderson, C. 2000. *Assessing Reading*. Cambridge: Cambridge University Press.
- Cook, P., Dixon, W., Duckworth, M., Kaiser, K., Koehler, W., Meeker & Stephenson, W. (2000). *Beyond Traditional Statistical Methods*. Iowa: Iowa State University Press.
- Foster, J. 2001. *Data Analysis Using SPSS for Windows*. London: Sage Publications Ltd.
- Fulcher, G. 1997. Text difficulty and accessibility: Reading Formulas and expert judgment. *System*, 25, 4, 497-513.
- Graesser, A., McNamara, D., Louwse, M. & Cai, Z. 2004. Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments & Computers*, 36, 2, 193-202.

# The production of Spanish vowels by early and late Spanish-English bilinguals

Natalia Mazzaro<sup>1</sup>, Laura Colantoni<sup>2</sup>, Alejandro Cuza<sup>3</sup>

<sup>1</sup>Department Languages & Linguistics, University of Texas at El Paso, U.S.

<sup>2</sup>Department of Spanish & Portuguese, University of Toronto, Canada.

<sup>3</sup>Department of Spanish & Portuguese, Purdue University, U.S.

## Abstract

We examine the production of the Spanish vowels [a,e,i] by two groups of Spanish-English bilinguals plus a control group from the same dialectal area. The bilinguals differed in their age of onset of acquisition of English; the early bilinguals acquired English before puberty and the late bilinguals after puberty. Participants read a list of words containing the target sounds controlled by position, stress and phonetic context. Formant values (F1 and F2) were measured at 3 points in each vowel, to investigate diphthongization. We hypothesized that early bilinguals' production of vowels would exhibit (i) fronting, (ii) diphthongization and (iii) reduction, following the English norm. Hypotheses (i) and (ii) were confirmed. We discuss the influence of age of arrival on bilinguals' performance.

Key words: Spanish-English bilingual, Spanish vowel, acoustic analysis, attrition.

## Introduction

Bi-directional influences have been documented at all linguistic levels, including phonetics and phonology (Antoniou *et al.* 2011), and Spanish-English bilinguals are not an exception (e.g. Bradlow 1995). The debate centers on whether maturational constraints and age of onset of acquisition (AOA) are responsible for such influences. If maturational approaches are correct (Abrahamsson & Hyltenstam 2009), early bilinguals (heritage speakers) are expected to behave as late bilinguals (long term immigrants), since both acquired Spanish before puberty. However, if AOA affects the perception and production of sounds, then early bilinguals exposed to English from birth, and with limited exposure to Spanish, will have difficulties producing native-like Spanish contrasts due to reduced Spanish input and intense contact with English.

If age of onset of bilingualism affects the production of sounds, we hypothesized that early bilinguals will: (i) produce front vowels with a more fronted tongue position (following the English norm) relative to the tongue position for the Spanish vowels (Bradlow 1995); (ii) Exhibit a higher degree of diphthongization. (iii) Have greater differences in duration between stressed and unstressed vowels. On the other hand, late bilinguals who acquired the L2 past puberty should behave closer to the attested monolingual patterns (MacKay *et al.* 2001).

## Method

To determine the relative role of maturational constraints and AOA, we studied two groups of bilinguals who spoke Spanish from birth but differed in their age of onset of acquisition of English. Whereas the early bilinguals were exposed to English during early childhood and received formal education in English, the late bilinguals were exposed to English after puberty and were educated in Spanish. A total of twenty-four participants (n=24) participated in the study: 12 early bilinguals (mean AOA=2.5; mean length of residence= 21.2), 6 late bilinguals (mean AOA=24.3; mean length of residence=14.5) from El Paso, Texas, and a group of 6 monolinguals from Chihuahua – Mexico.

We examined the production of Spanish vowels [a], [e], [i] preceded by either voiceless or voiced stops in stressed and unstressed syllables. Subjects read real words (n=112) in a carrier phrase *Digo X para ti* ('I say X for you').

The acoustic dimensions analysed were vowel formants (normalized using the Nearey 1 method), measured at different time points (25, 50 and 75%) into the vowel to detect diphthongization, and relative duration of the unstressed vs. stressed vowels.

The degree of diphthongization was measured by using the formula:

$$D = \sqrt{(F1_{25} - F1_{75})^2 + (F2_{25} - F2_{75})^2}$$

Where  $F1_{25}$  and  $F2_{25}$  are the values of F1 and F2 measured 25% into the vowel, and  $F1_{75}$  and  $F2_{75}$  are the values of F1 and F2 measured 75% into the vowel.

The different groups were compared using one-way ANOVAs.

## Results

Hypothesis i) was confirmed for /i/. Early bilinguals have a significantly higher normalized F2 (thus more fronted /i/) than late bilingual and monolingual speakers (Table 1 and Figure 1). This effect was not found in /a/ or /e/, where groups are clustered together. However, the lack of variation in /a/ and /e/ should be taken with caution, since the normalization method could have erased important variation in these vowels.

Hypothesis (ii) was confirmed for /a/ and /i/. Early bilinguals more frequently diphthongize /a/, while late bilinguals do the same with /i/ (Table 2). Stress does not seem to affect this trend.

Table 1. Normalized F2 values for [a], [e] and [i] measured at vowel midpoint.

	[a]		[e]		[i]	
	mean	SD	mean	SD	Mean	SD
Early biling.	0.93	0.06	1.07	0.06	1.16	0.11
Late biling.	0.94	0.05	1.07	0.07	1.14	0.13
Monolingua	0.94	0.06	1.07	0.06	1.12	0.10
<i>p</i> -value	0.001		0.82		<0.001	

Table 2. Change in normalized F1-F2 distance from 25% to 75% of the vowel duration.

	[a]		[e]		[i]	
	mean	SD	mean	SD	Mean	SD
Early biling.	1.23	1.12	1.05	1.24	1.29	1.53
Late biling.	1.17	1.03	1.04	1.02	1.82	2.09
Monolingua	1.06	0.85	0.83	0.64	1.00	1.08
<i>p</i> -value	0.03		0.07		<0.001	

Hypothesis (iii) was rejected. Early bilinguals' relative duration of unstressed vs stressed vowels is similar to those of monolinguals. Yet, late bilinguals produce significantly larger differences between stressed and unstressed vowels.

Our results suggest that reduced Spanish input and intense contact with English affect the production of vowels. Bilingual speakers produced /i/ with a more fronted articulation and /e/ and /a/ with a higher degree of diphthongization compared to monolinguals. It is possible that bilinguals have established a common phonetic category for Spanish /i/ and English /i/, which is shown by its shift towards the front. However, since /e/ and /a/ could be mapped to different English vowels (/e, ɪ, ε/ and /æ, ɒ, ʌ, ε/, respectively) (Morrison 2003), it is possible that new phonetic categories were created for the English vowels, leaving the Spanish ones unaffected. Late bilinguals' reduced difference in the duration of stressed and unstressed vowels compared to monolinguals and early bilinguals was unexpected. Yet, research has shown that increased use of L2 is associated with inhibition of

L1 (Linck et al., 2009), which may result in movement away from the native norm.

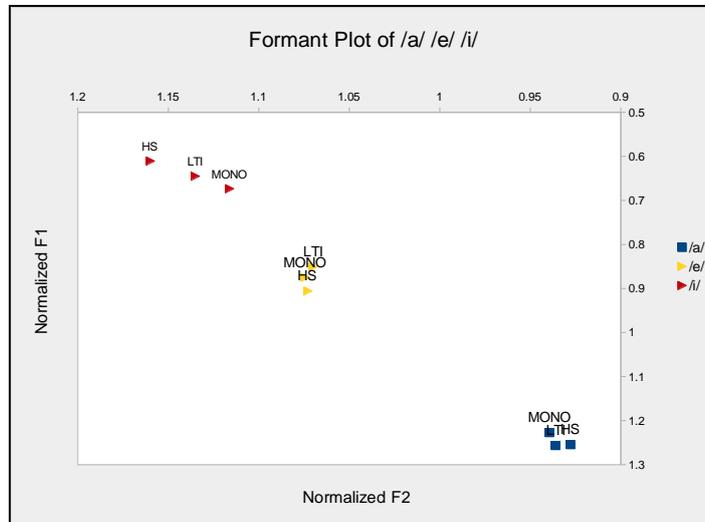


Figure 1. Scatterplot of /a, e, i/ for all three groups.

## Acknowledgements

We wish to thank Martin Lazzari for his support with the statistical analysis, and the members of the Language Acquisition & Linguistics Research Laboratory at UTEP for their help with the acoustic analysis of the data.

## References

- Abrahamsson, N. & Hyltenstam, K. 2009. Age of onset and native-likeness in a second language: listener perception versus linguistic scrutiny. *Language Learning*, 59:2, 249–306.
- Antoniou, M., Best C., Tyler M., Kroos, C. 2011. Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics* 39, 558–570
- Bradlow, A. 1995. A comparative acoustic study of Spanish and English vowels. *Journal of the Acoustical Society of America*, 97, 1916-1924.
- Linck, J., Kroll, J. & Sunderman G. 2009. Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychological Science* Volume 20 (12), 1507-1515.
- MacKay, I. R. A., Flege, J. E., Piske, T., & Schirru, C. 2001. Category restructuring during second-language speech acquisition. *The Journal of the Acoustical Society of America*, 110(1), 516-528.

Morrison, G. 2003. Perception and production of Spanish vowels by English speakers. In Solé, MJ, Recansens D. & Romero J. (eds.), Proceedings of the 15th International Congress of Phonetic Sciences: Barcelona 2003. 1533–1536. Adelaide, South Australia: Causal Productions.

# Eye movements reflect acoustic cue informativity and statistical noise

Jessie S. Nixon<sup>1,2,3</sup>, Jacolien van Rij<sup>4</sup>, Peggy Mok<sup>5</sup>, Harald Baayen<sup>4</sup>, Yiya Chen<sup>1,2</sup>

<sup>1</sup>Leiden Institute for Brain and Cognition (LIBC), Leiden University, Netherlands

<sup>2</sup>Leiden University Centre for Linguistics, Leiden University, Netherlands

<sup>3</sup>MARCS Institute, University of Western Sydney, Australia

<sup>4</sup>Department of Linguistics, University of Tübingen, Germany

<sup>5</sup>Department of Linguistics, Chinese University of Hong Kong, Hong Kong

## Abstract

Listeners rely on highly variable, non-discrete acoustic information to understand spoken messages. The present ‘visual world’ eye tracking study investigated whether the *amount of acoustic cue variation* affected Cantonese listeners’ perception of speech contrasts. Participants saw pictures of word pairs which were identical except for initial consonants (unaspirated versus aspirated). Auditory stimuli were continua of increasing VOT presented in bimodal distributions. The amount of acoustic variation varied between conditions: high-variance versus low-variance. *Generalised Additive Modelling* analyses showed, in the low-variance condition, eye movements reflected cue values: there was differential fixation behaviour for category means, boundaries and peripheries. In contrast, in the high-variance condition, the acoustic cue had little effect: fixation behaviour was similar across the different acoustic cue values. This demonstrates listeners’ high sensitivity to the discriminative value of acoustic cues. How much cue dimensions are utilised depends on their variance.

Key words: speech perception, acoustic variation, word discrimination, Cantonese

## Introduction

The acoustic information that listeners use to discriminate between possible spoken messages is highly variable and non-discrete. Speech consists of a complex collection of cues, which listeners can potentially use for discrimination. Which factors affect whether and the extent to which listeners use a particular cue? There is substantial evidence that infants and adults detect and respond to the *number of peaks* in a distribution, i.e. unimodal versus bimodal distributions (e.g. Maye et al., 2002). However, little is known about how *acoustic variation* affects speech perception. In one innovative recent study to address this question, Clayards et al. (2008), showed that increased within-category variability led to a greater proportion of looks to competitor objects in English stop discrimination. However, only cue values near distribution peaks were analysed and fixation proportions were averaged over trials, so there was no analysis of the time course.

The present study investigates how acoustic cue variability affects native Cantonese listeners' discrimination during speech perception. This informs the question of how listeners are able to utilise informative cues and ignore irrelevant ones. Importantly, the present study also investigates the *time course* of effects. We analysed fixations over the course of the trial. In addition, we also analysed the whole *cue continuum*, to examine the effects on cue values near the category boundary and perimeters. We use the term *variance* to describe the amount of *within-category* acoustic variation. This refers to the degree to which acoustic values spread out from the category mean. A variance of zero means that all values are identical.

## Method

### Participants

Thirty-seven native Cantonese-speaking students from the Chinese University of Hong Kong participated in the experiment for payment.

### Experiment design and stimuli

The experiment design and stimuli were based on Clayards, et al. (2008). Visual stimuli were picture pairs, identical except for initial consonants, which were *unaspirated* (bou3, `cloth'; jun1 `brick') or *aspirated* (pou3, `shop'; chun1, `village'). Auditory stimuli were recorded by a native Cantonese speaker and resynthesised into a 12-step VOT continuum. All participants heard a bimodal distribution of auditory stimuli. Only token presentation frequency varied between conditions: high-variance vs. low-variance.

### Procedure

Participants wore an SR Eyelink II eye-tracker. The session started with familiarisation, then a practice. Trials consisted of a brief (1000 ms) preview of four pictures, followed by a gaze-contingent fixation cross, then auditory stimuli were played simultaneously as the pictures reappeared. Eye movements were monitored until participants clicked on the picture they heard.

### Analysis and results

Eye movement data were analysed using *Generalized Additive Modelling* (GAM; Wood, 2006, 2011) using the *mgcv* package in R. GAMs are a type of regression analysis that drop the linearity assumption. The degree of non-linearity is determined from the data itself: the optimal linear or non-linear equation to avoid model over-fitting and over-generalizing (Wood, 2006).

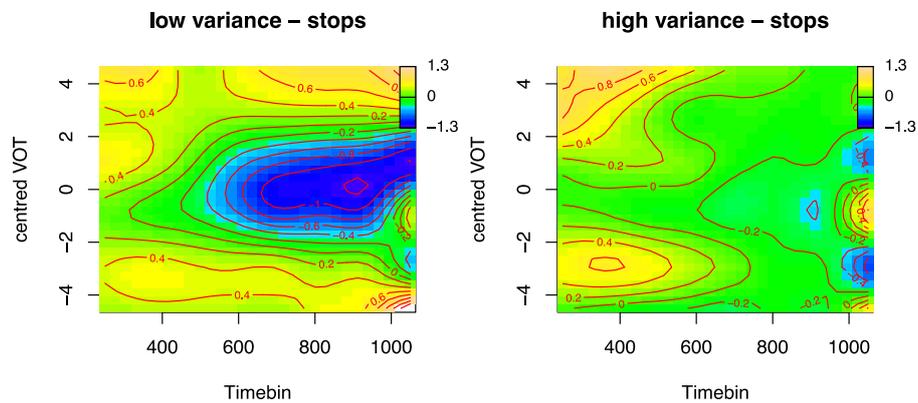


Figure 1. Topographical maps of the proportion of fixations on the clicked target versus competitor for VOT over time in the GAM model for stops in the low-variance (left panel) and high-variance conditions (right panel). Estimated effects are on logit scale. Time (x-axis) is in bins of 100 ms. VOT (y-axis) is centred around 0, the category boundary. Negative VOT values correspond to unaspirated stimuli, positive values to aspirated stimuli. Category means are at VOT -2.5 and 2.5. Fixation proportions (z-axis) are represented by colour codes. Positive values (yellow) indicate relatively more looks to the clicked target; negative values (blue) indicate relatively more looks to the competitor. Random effects are excluded from these plots.

Model comparisons, model summaries and model plots all provide evidence for an effect of distribution condition. A  $\chi^2$  test of fREML scores which takes into account model complexity showed that a VOT by condition interaction over time significantly improved model fit ( $\chi^2(3)=135.32$ ,  $p<.001$ ). Since this interaction was significant, lower-level predictors were retained in the model. The model also included manner of articulation (stops versus affricates), but due to space limitations, we will focus on our main predictor of interest, distribution condition. The model summary showed that the variance explained by surfaces of VOT over time ( $\chi^2(32.95)=3637.4$ ,  $p<.001$ ) and VOT by condition over time ( $\chi^2(28.27)=3533.11$ ,  $p<.001$ ) were significantly different to zero.

The model plots (Figure 1) show differential patterns of looking behaviour between the low-variance (left panel) and high-variance conditions (right panel). The low-variance condition displays divergent patterns between category means, category boundaries and the distribution peripheries. The high-variance condition, in contrast, is quite flat across VOT values after about 500 ms after presentation of the auditory stimulus.

## Discussion

The present study demonstrates that subtle differences in acoustic cue variance can have immediate effects on the way a particular cue is perceived. The GAM analysis revealed that fixations on the clicked target over time were affected by VOT value. More interestingly, the effect of VOT significantly interacted with distribution condition. From about 500 ms into the trial, distinct eye movement patterns emerged for different VOT values in the low-variance condition (left panel, Figure 1), with differential fixation behaviour at category means, boundaries and distribution peripheries. In contrast, in the high-variance condition (right panel, Figure 1), VOT had a weaker effect: fixation behaviour was similar for all VOT values. The eye movement patterns seem to reflect different stages of processing: an initial, perceptual stage and a later process of verification. This suggests that, at least at later stages of processing, the acoustic cue is relied on less for discrimination in the high-variance condition, when it is less informative, than in the low-variance condition when it is more informative.

In summary, the present results demonstrate that the discriminative value of acoustic cues has an immediate effect on how they are processed in speech perception. When cues more consistently fall within a small range of values, they are more reliable as discriminators, and consequently used more effectively for discrimination. However, when cues are highly variable over a range of values, their effectiveness for discrimination declines, and the degree to which they are utilised in perception decreases in turn. Interestingly, the analysis of the time course shows that this is a relatively late effect. This may suggest that, following initial perceptual processes, there is increased difficulty at a later verification stage.

## Acknowledgements

JSN was responsible for the experiment design, implementation, stimuli creation, data collection and manuscript preparation. JSN and JvR conducted the analysis and interpretation with significant contributions from HB. PM assisted with finding equipment and participants. This work was supported by European Research Council ERC Starting [Grant 206198] to YC.

## References

- Clayards, M., Tanenhaus, M., Aslin, R., Jacobs, R. A., 2008. Perception of speech reflects optimal use of probabilistic speech cues. *Cognition* 108, 804–809.
- Maye, J., Werker, J.F., Gerken, L., 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82 (3).
- Wood, S. 2006. *Generalized additive models: an introduction with R*. CRC press.

# Cross-category phonological effects on ERP amplitude demonstrate context-specific processing during reading aloud

Jessie S. Nixon<sup>1,2,3</sup>, Jacolien van Rij<sup>4</sup>, Xiaoqing Li<sup>5</sup>, Yiya Chen<sup>1,2</sup>

<sup>1</sup>Leiden Institute for Brain and Cognition (LIBC), Leiden University, Netherlands

<sup>2</sup>Leiden University Centre for Linguistics, Leiden University, Netherlands

<sup>3</sup>MARCS Institute, University of Western Sydney, Australia

<sup>4</sup>Institute of Psychology, Chinese Academy of Science, China

<sup>5</sup>Department of Linguistics, University of Tübingen, Germany

## Abstract

The phonetic realisation of speech sounds depends on their context: e.g., /t/ is aspirated in ‘top’, but unaspirated in ‘stop’. Similarly, Beijing Mandarin Tone 3 (T3) usually has a low contour, but preceding another T3 syllable has a rising contour (sandhi T3). Importantly, Tone 2 (T2) also has a similar rising contour. Yet how such phonetic variation is processed online during reading and speech production is not well understood. ERP amplitude was measured as native Mandarin speakers read aloud words preceded by a 48ms masked prime. Prime and target always differed in tone category. Critical targets were T2-initial words. Primes were either sandhi T3 (*contour-match*) or low T3 (*mismatch*). Generalised additive mixed models (GAMs) revealed a complex interaction between prime type, prime frequency and target frequency emerging around 100ms and 300-400ms following target presentation. In the contour-match condition, when prime and target frequency were both high or both low, there was increased negativity, suggesting competition between prime and target. In the mismatch condition, there was relatively little effect of item frequencies. This difference in the pattern of effects between the contour-match and mismatch primes provides evidence for top-down effects of context on phonological processing of masked primes during reading aloud.

Key words: reading aloud, context effects. EEG / ERPs, Mandarin Chinese, masked priming, tone processing, allophonic variation.

## Introduction

Phonetic variation is a fundamental property of speech. Speech sounds vary depending on their context. For example, in words like ‘spin’, where the first sound is /s/, the second sound is considered to belong to the same sound category /p/ as in the word ‘pin’. But acoustically, the voice onset time falls between the /p/ of ‘pin’ and /b/ of ‘bin’. How this kind of ‘within-category’ variation is processed is not yet well understood.

In Beijing Mandarin, Tone 3 (T3) has at least two variants. T3 usually has a low contour, but preceding another T3 syllable, it is realised with a rising contour (sandhi T3). Importantly for the present study, the contour of

sandhi T3 is very similar to another tone, Tone 2. A previous study (Nixon, Chen & Schiller, 2015) showed that production and visual processing of Mandarin tonal variants involves multi-level phonological processing, where there is activation of both the tonal contour and the tone category. The present study investigates whether the context provided by the following character in briefly presented masked primes affects processing of tone during reading aloud. In addition, it investigates whether phonological similarity (contour overlap) between prime and target can affect processing even when prime and target differ in tone category.

## Method

### Participants

Twenty-four native speakers of Beijing Mandarin participated in the experiment for pay.

### Design and stimuli

Critical targets consisted of 25 two-character Chinese words, of which the initial character was Tone 2 (e.g. 鱼缸, *yu2gang1*, 'fish tank'). Each target was preceded by a two-character prime. The initial syllable of each prime was a Tone 3 character, which had the same segmental syllable as the first character of the target. Therefore, all prime-target pairs differed in terms of the tone category. The second character was either T3 (i.e. sandhi word, *contour-match prime*, e.g. 雨水, *yu3shui3*, 'rain') or a different tone (i.e. low-tone word, *mismatch prime*, e.g. 雨衣, *yu3yi1* 'raincoat'). For each target word, the initial character of the prime was identical between prime conditions. They either matched or mismatched in contour.

### Procedure

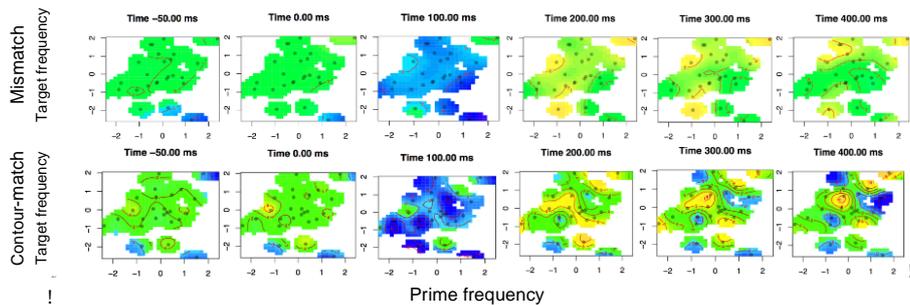
Stimulus presentation and reaction time data acquisition were conducted using the E-Prime 2.0 with a voice key trigger. The session began with a practice block to familiarise participants with the procedure. Each experimental trial began with a fixation cross with jittered presentation time (400-700 ms) to reduce time-induced expectancy waves. A forward-mask followed for 100 ms, before presentation of the prime for 48 ms. A backward mask was presented for 17 ms to avoid images of the prime remaining on the retina. Finally, the target word was presented for a maximum of 2000 ms or until the participant response, which triggered the voice key and caused the word to disappear. The experimenter coded incorrect responses and voice key errors in a 1400 ms interval before the beginning of the next trial. Response

time was calculated from the time of target word presentation until the voice key was triggered by the participant response.

## Analysis and Results

ERPs were analysed using *Generalised Additive Mixed Modelling* (Wood, 2006), a non-linear model with random effects for subjects and items. Analysis was conducted on a set of 24 electrodes evenly distributed over the scalp. To avoid potential speech artefacts, only data preceding 500ms after target presentation were analysed. For 19 of the 24 electrodes, model comparisons and model summaries showed a significant interaction between prime type, prime frequency and target frequency over time. Figure 1 shows the interaction between prime frequency and target frequency over time for the mismatch (top row) and contour-match conditions (bottom row) at electrode Fz. Each panel shows the interaction between Johnson transformed prime frequency and target frequency. ERP amplitude is colour-coded.

ERP amplitude varies between prime types (contour-match vs. mismatch) and interacts with prime and target word frequency. Divergences emerge around 100ms and 300-400ms after target presentation. In the contour-match condition, negativity is reduced when target frequency is relatively high and prime frequency low. However, when prime and target frequency are both high or both low, there is increased negativity, suggesting competition between prime and target. In the mismatch condition, there is relatively little effect of item frequencies.



**Figure 1.** Topographical map of ERP amplitude at Fz. the interaction between prime frequency and target frequency over a series of time points (-50ms, 0 ms, 100 ms, 200 ms, 300 ms, 400 ms) for trials with the mismatch prime (top row) and contour match prime (bottom row) for electrode Fz. Each panel shows the interaction between Johnson transformed prime frequency (horizontal axis) and Johnson transformed target frequency (vertical axis). Yellow indicates relatively more positive amplitude; blue indicates more negative amplitude; green is at the intercept. The grey dots indicate the individual items.

## Discussion

The present results provide evidence for automatic retrieval of sub-phonemic information in visually processed masked prime words. The context provided by the tone of the second character of the prime determined whether the prime was rising, and therefore matched the target, or low, which mismatched the target. Model comparisons, model summaries and plots all showed significant differences in the pattern of ERPs when prime and target matched in tonal contour, compared to when they did not match. This demonstrates that even for very briefly presented masked primes, the context-specific phonetic form is processed. Moreover, activation of this context-specific variant influences processing of targets from a different tone category, demonstrating cross-category phonological effects.

The analysis shows a complex interaction between prime type, prime frequency and target frequency. In the contour-match condition, the item characteristics came into play. When both prime and target frequency were high, there was increased negativity, suggesting an increased processing cost of resolving competition between activated candidates. In the mismatch condition, in contrast, there was little effect of the item frequencies. This last finding raises a methodological issue. In priming studies, interactions between prime and target frequency can have significant effects on ERP amplitude. Therefore, in ERP studies of language, it may be important to include item characteristics in the statistical analysis.

## Acknowledgements

This work was supported by the European Research Council under the ERC Starting [Grant 206198] to YC.

## References

- Nixon, J., Chen, Y. and Schiller, N. (2015). Multi-level processing of phonetic variants in speech production and visual word processing: evidence from Mandarin lexical tones. *Language, Cognition & Neuroscience*. 30 (5)491-505. DOI: 10.1080/23273798.2014.942326
- Wood, S. (2006). Generalized additive models: an introduction with R. CRC press.

# Realization of Serbian accents by Serbian and Russian speakers

Ekaterina Panova

Department of history and theory of language, St. Tikhon's Orthodox University,  
Russia

## Abstract

The present study examined realization of Serbian pitch accents by two Serbian and four Russian female speakers in initial, medial and final position of the statements. For each word the set of pitch parameters were calculated. The main statistical differences between realizations of rising and falling accents by Serbian speakers were found with respect to F0 start value and timing of F0 maximum. These pitch parameters also served as an effective indicator of deviations which were observed across Russian speakers. The main difficulties for Russian speakers were concerned realizations of falling accents in initial and medial positions.

Key words: pitch contour, pitch accent, Russian, Serbian.

## Introduction

One of the main differences between Russian and Serbian stress is a tone character of the latter. Traditionally, it has been established that standard Serbian has four accents, which differ in pitch (rising or falling) and duration (short or long): long rising (LR), long falling (LF), short rising (SR) and short falling (SF). The problem of the distinction between falling and rising accents is discussed in many investigations (for example, in Leniste 1986, Keijsper 1987). The most valuable parameters of the distinction are range of F0 between accented and post-tonic syllable and location of the F0 peak.

Russian stress is mostly characterized by duration and vowel quality. In Russian F0 is the main parameter of intonation, since Russian belongs to the languages with strong influence of intonation on word prosody (Nikolaeva 1977). In Serbian intonation also has a great influence on word prosody, although types of accents can modify its contour. In final position of the statements there is a tendency in neutralization of the accents (Lehiste et al. 1986).

Despite the fact that within the last decade many phonetic studies of Serbian and Russian have appeared, there are no comparisons of prosodic features between these languages except one extensive study, which was conducted in 1970's by Nikolaeva (Nikolaeva 1977). The work presented here forms part of a larger study which aims to investigate how Russian speakers realize Serbian accents and develop the methodology for training Serbian accents by Russian speakers.

## Method

For the present study 111 target words of mono-, di- and tri-syllables with different vowel inventory and different types of accent were selected from available dictionaries. For minimizing segmental influences only words with unvoiced consonants were selected. Each target word was embedded in frame sentences so as to occur in initial, medial and final position. All the sentences were statements.

Then, two control native speakers of Serbian (S1, S2, females) and four Russian speakers (R1, R2, R3, R4, females) read these sentences in neutral style and normal tempo.

For both Serbian and Russian sentences the contour of F0 was calculated. Visual and auditory analyses were conducted to eliminate all misreadings and errors in calculation of F0. From contours of the target words the set of pitch parameters (described in detail in (Smirnova et al. 2007) were obtained and then processed with the STATISTICA software program.

## Results

Here the results for two pitch parameters, F0 start value and timing of F0 maximum are presented.

### F0 start value

A two-way ANOVA with repeated measures (type of accent x position) was performed for Serbian and Russian speakers separately. For this statistical analysis four types of Serbian accents were divided only in two types: falling (FA) and rising (RA).

Post hoc tests Tukey HSD indicated that there was a significant difference between falling and rising accents for Serbian speakers in initial ( $p < 0.0001$ ) and medial positions ( $p < 0.0001$ ) and only for one Serbian speaker (S1  $p = 0.226$ , S2  $p < 0.0001$ ) in final position. F0 start value serves as an indicator of the types of accents: rising accents have lower F0 start value, than falling ones. This difference between accents can be observed even in final position (i.e. nucleus), which usually influences much on word prosody.

Across Russian speakers there was no significant difference between the types of accents ( $p = 1.000$ ) except for one Russian speaker, who has a significant difference only in medial position (R3  $p = 0.001$ ). Although some Russian speakers may have the same tendency as Serbian speakers, they don't demonstrate any significant dependence (Figure 1).

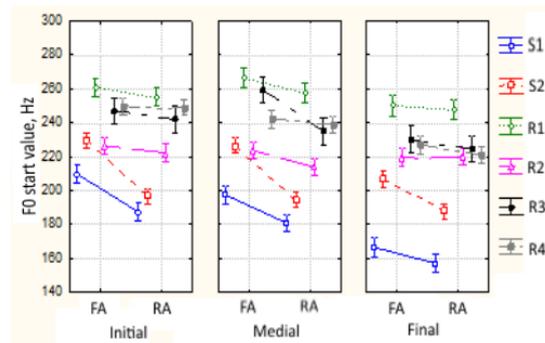


Figure 1. F0 start value scores of the words with falling (FA) and rising accents (RA) for Serbian (S1, S2) and Russian speakers (R1, R2, R3, R4) in initial, medial and final position.

### Timing of F0 maximum

For the estimation of timing of F0 maximum the Survival Analysis was performed with type of accent as the grouping variable. For Serbian speakers there was a significant difference between accents in initial and medial position ( $p < 0.0001$ ), but not in final position ( $p > 0.1$ ). Thus, falling accents mostly have earlier peak locations (up to 40%) and rising accent have later peak locations (from 50%). In final position the distinction between rising and falling accents is neutralized (Figure 2).

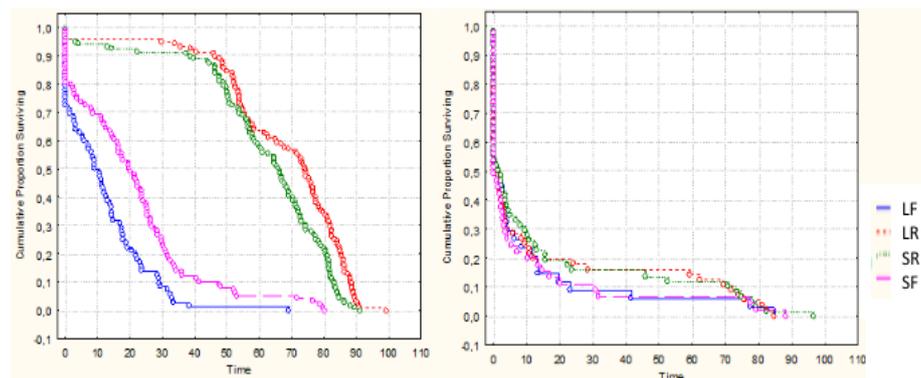


Figure 2. Timing of F0 maximum scores of words with falling (FA) and rising accents (RA) in initial, medial (left) and final position (right) for Serbian speakers.

Russian speakers didn't have strong tendency between the types of accents in timing of F0 maximum in all positions ( $p > 0.1$ ). The values of

timing of F0 maximum showed that Russian speakers mostly realized “rising” accents instead of falling ones (from 50%, Figure 3).

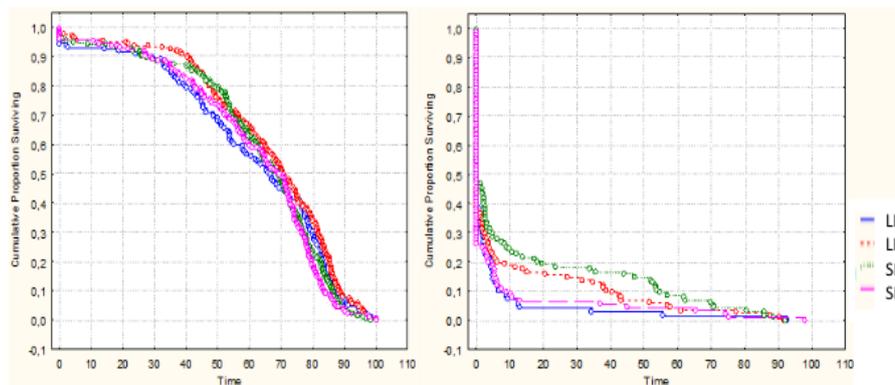


Figure 3. Timing of F0 maximum scores of words with falling (LF, SF) and rising accents (LR, SR) in initial, medial (left) and final position (right) for Russian speakers.

## Conclusion

The results of the present study showed that pitch parameters, F0 star value and timing of F0 maximum, allowed to provide the distinction between Serbian rising and falling accents, and served as an effective indicator of deviations which were observed across Russian speakers.

Although previous researches of Serbian accentuation mentioned parameters of F0 star value (Lehiste 1981), it isn't interpreted as a one of the most distinctive between rising and falling accents.

The main difficulties in realizations of Serbian accents by Russian speakers are related to falling accents in initial and medial position. In final position the neutralization of the accents is observed across both Serbian and Russian speakers.

## References

- Keijsper, C.E 1987. Studing Neoštokavian Serbocroatian Prosody. Dutch Studies in South Slavic and Balkan Linguistics. SSGI 10, 101-193.
- Lehiste I., Ivic P. 1986. Word and sentence prosody in Serbocroatian. Cambridge, Mass., MIT Press.
- Nikolaeva T.M. 1977. Frazovaja intonacija slavianskih jazykov. Moskva, Nauka.
- Smirnova, N., Starshinov A., Oparin I. & Goloshchapova T. 2007. Speaker Identification Using selective Comparision of Pitch Contour Parameters. Proc. of the 16<sup>th</sup> ICPhS, Saarbrucken, 203–206.

# Effects of normal aging on obstruents' acoustic characteristics

Eirini Papadiamantopoulou<sup>1</sup>, Elina Nirgianaki<sup>2</sup>

<sup>1</sup>Department of Speech & Language Therapy, Akmi Metropolitan College, Greece

<sup>2</sup>Lab of Phonetics & Computational Linguistics, University of Athens, Greece

## Abstract

The present study constitutes an experimental investigation on adult female's temporal and spectral (formant transitions) parameters of Greek voiceless stop consonants /p, t, k/ and fricative consonant /s/, as functions of normal aging. The effects of place of articulation and syllabic structure on these acoustic parameters were also examined. Results indicated that the two age groups were only distinguished by the voice onset time (VOT) duration of the velar voiceless stop /k/ and by the F2 transition values after /p/ and /t/. A main effect of aging on fricative /s/ duration was not confirmed. In addition, VOT distinguished stops' place of articulation, and the factor of syllable structure was significant for /s/ duration.

Key words: aging, duration, formant transitions, Greek, voiceless obstruents

## Introduction

The aging process causes anatomical and physiological changes in the speech production system, including laryngeal, respiratory, articulatory and resonance mechanisms (Vipperla, Renals and Frankel 2010).

Several studies have shown that the aging factor has a significant effect on formant frequencies (Xue and Xao 2003; Das et al 2012; Sfakianaki, n.d.). Additionally, it has been revealed that older speakers produce shorter VOTs when compared to younger groups (Torre and Barlow 2005; Yao, 2007). Aging has also been found to affect segmental duration in a proportional manner (Smith, Wasowicz and Preston, 1987; Brazeal, 1997; Nissen and Fox 2005; Yao 2007).

However, all these studies either examine the acoustic characteristics of sounds in other languages than Greek (e.g. Xue and Xao, 2003; Das et al 2012) or concern the comparison of sounds produced by children and those produced by adults (e.g. Sfakianaki, n.d.).

In this frame, the present study aims at investigating the aging effect on Greek adults in 1. VOT duration of voiceless stops, 2. fricative /s/ duration, and 3. post consonantal formant frequencies (F1, F2) after stops. Simultaneously, it examines: 1. intragroup differences between VOT duration of stops, 2. the effect of syllable structure (/s, sk, ks/) on /s/ duration, and 3. intragroup differences of F1, F2 values after stops.

## Methodology

Ten female speakers were equally divided into two age groups: Group A, aged between 20-30, and Group B, aged between 60-70. Two-syllable words of CVCV and CCVCV structure, stressed on the first syllable were produced by both groups in three repetitions. The Greek words /'para, 'tara, 'kara, 'sara, 'skara, 'ksara (nonsense)/were placed in the carrier sentence /i 'leksis ... 'ine elini'ci/ (The word ...is Greek).

Acoustic analysis was carried out in PRAAT (version 5.3.77). Both stops' Voice Onset Time (VOT) and fricative's duration measurements (in ms) involved the simultaneous consultation of waveform and wideband spectrogram. For the stop-vowel transitions, the formant loci for F1 and F2 (Hz) were identified in reference to the spectrographic display as follows: 1. By the Praat-derived Formant Tracker, and 2. By numerical cursor readouts on the screen. Both F1 and F2 were calculated at vowel onset, at the first glottal pulse following stop offset. Statistical analysis was carried out in SPSS 17.0.

## Results

Mean VOT durations of voiceless stops differed between the two groups, but t-tests revealed significant difference only for /k/ [ $t(28)=-4.449$ ,  $p<0.05$ ], with older females producing longer VOT than younger ones (Fig.1). Moreover, place of articulation had a main effect on VOTs produced by the young group [ $F(44)=85.663$ ,  $p<0.0001$ ] with significant differences between /p/-/k/ and /t/-/k/; and by the older one [ $F(44)=122.763$ ,  $p<0.0001$ ] with significant differences among all three places (Fig. 1).

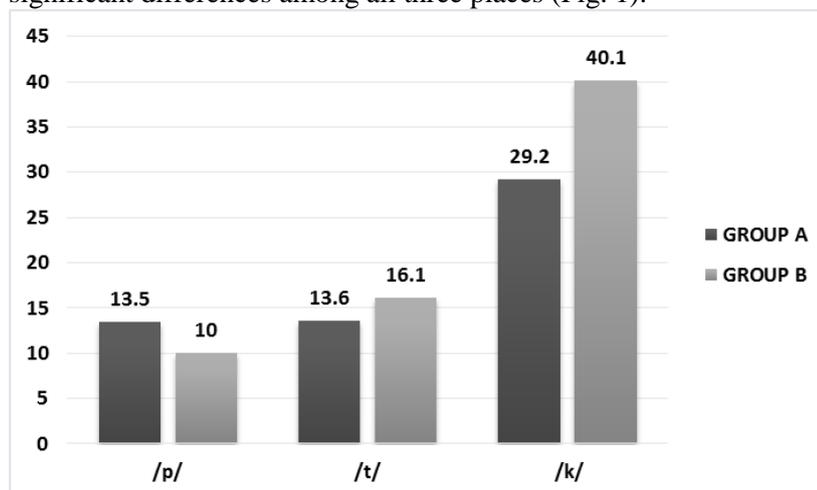


Fig. 1. Mean VOT duration (ms) for /p/, /t/, and /k/ as a factor of aging.

Regarding /s/ duration, aging had no significant effect on it [ $t(28)=4.273$ ,  $p>0.05$ ]. Nonetheless, one way ANOVA showed that the syllable structure interacted significantly with fricative duration [ $F(89)=48.431$ ,  $p<0.0001$ ]. Absolute mean duration of /s/ for all speakers was longer when the sound was produced as singleton (142 ms), followed by /s/ in /ks/ (110 ms) and /s/ in /sk/ (107 ms).

Concerning formant transitions, there were no significant differences between age groups in F1. However, t-tests revealed significant differences in F2 after /p/ [ $t(28)=0.798$ ,  $p<0.05$ ], and /t/ [ $t(28)=2.397$ ,  $p<0.05$ ], with the younger group exhibiting higher values, but no significant differences after /k/ [ $t(28)=1.899$ ,  $p>0.05$ ] (Table 1).

Place of articulation affected significantly F1 for both the young [ $F(44)=5.587$ ,  $p<0.05$ ] and the older [ $F(44)=5.274$ ,  $p<0.05$ ] group. Post hoc tests revealed significant differences between F1 after /p/ and /t/ ( $p<0.05$ ) for the young group, and /p/ and /k/ for both groups ( $p<0.05$ ), with F1 being highest after /p/. Similarly, significant differences were revealed for both groups [Group A:  $F(44)=77.883$ ,  $p<0.0001$ ; Group B:  $F(44)=16.105$ ,  $p<0.0001$ ] between F2 after /p/ and /t/, and /p/ and /k/ ( $p<0.0001$ ), with F2 being lowest after /p/ (Table 1).

Table 1. Mean F1 and F2 values (Hz) for both age groups (A and B) after stop consonants.

Consonant	F1		F2	
	A	B	A	B
/p/	770	785	1445	1398
/t/	658	665	1869	1764
/k/	653	614	1844	1721

## Discussion

To sum up, aging was revealed to affect significantly VOT duration of only the velar voiceless stop /k/, with older females producing longer VOT (40 ms) than younger ones (29 ms). These results agree with Brazeal (1997), who reported that older speakers produce longer VOT for /k/, and similar VOT for /t/ than younger ones, but disagree with Torre and Barlow (2005), and Yao (2007), who report shorter VOTs for the older participants. However, the present results show that the age factor does not affect the fricative /s/ duration, although Nissen and Fox (2005), who examined /s/'s durational differences between children and adults, revealed shorter durations for children.

Our results confirm the fact that VOT provides classification cues for voiceless stops' place of articulation, in accordance with Fourakis (1986), who reported long duration for /k/, medium for /t/ and short duration for /p/. It also confirms that /s/ as a singleton has shorter duration than in clusters (/sk/ /ks/), also reported by Botinis et al (1999).

With regards to formant transitions, both groups were found to exhibit similar mean F1 values after all voiceless stops, contrary to other research findings (Das et al 2012; Sfakianaki, n.d.). On the other hand, the age affected significantly F2 for bilabial and alveolar places of articulation, with the younger group exhibiting higher F2 frequencies when compared to the older group, in accordance with Das et al (2012), Sfakianaki (n.d.), and Xue and Xao (2003) who all have reported that F2 values decrease with aging.

In conclusion, it seems that there is a compensatory effect of aging on stop voiceless consonants, i.e. the bilabials and alveolars are distinguished by F2 onset, while velars by VOT duration.

## References

- Botinis, A., Fourakis and M., Prinou, I., 1999. *Prosodic effects on segmental durations in Greek*. Communication and Technology, EUROSPEECH 99, vol. 6, 2475-78, Budapest, Hungary.
- Brazeal, V., 1997. The effect of aging on elderly Female voice onset time and segment duration. DEd. Oklahoma State University.
- Das, B., Mandal, S., Mitra, P., Basu, A., 2012. Effect of aging on speech features and phoneme recognition: a study on Bengali voicing vowels. *Int J Speech Technol*. Springer Science+Business Media, LLC 2012.
- Fourakis, M.. 1986. A Timing model for word-initial CV syllables in Modern Greek. *J Acoust Soc Am* 79, 1982-1986.
- Nissen, S.L., Fox, R.A., 2005. Acoustic and spectral characteristics of young children's fricative productions:A developmental perspective. *J. Acoust. Soc. Am* 118(4).
- Sfakianaki, N.D. Acoustic characteristics of Greek vowels produced by adults and children.
- Smith., B.L., Wasowicz, J. and Preston, J., 1987. Temporal characteristics of the speech of normal elderly adults. *J Speech Lang Hear Res* (Abst).
- Torre, P., Barlow, J.A., 2005. Age-related changes in acoustic characteristics of adult speech. *J Commun Disord* 42, pp. 324-333.
- Vipperla, R., Renals, S. and Frankel, J., 2010. Ageing voices: The effect of changes in voice parameters on ASR performance. *J Audio Speech Music Process*.
- Xue, S.A. and Hao, G.J., 2003. Changes in the human vocal tract due to aging and the acoustic correlates of speech production: a pilot study. *J Speech Lang Hear Res* (Abst).
- Yao, Y., 2007. Closure duration and VOT of word-initial voiceless plosives in English in spontaneous connected speech, UC Berkeley Phonology Lab Annual Report.

# Effect of proficiency on subject-verb agreement processing in French learners of English: An ERP study

Maud Péliissier<sup>1</sup>, Jennifer Krzonowski<sup>2</sup>, Emmanuel Ferragne<sup>1</sup>

<sup>1</sup>Laboratoire CLILLAC-ARP, EA 3967, Université Paris Diderot, France

<sup>2</sup>Laboratoire DDL, UMR 5596, CNRS – Université Lyon 2, France

## Abstract

This study explored the effect of proficiency on the native-likeness of syntactic processing in a second language. ERP responses to violations of third-person singular subject-verb agreement were obtained from advanced and intermediate adult French learners of English as well as English native speakers. Our findings show a proficiency effect only in early ERP responses to violations, and suggest that the highest level of proficiency attainable through instruction in France is not enough to exhibit entirely native-like responses.

Key words: ERP, L2 processing, syntax, proficiency

## Introduction

Learning a second language (L2) as an adult is a difficult task, which rarely ends in the complete mastery of the target language. One of the factors identified to explain the remaining differences in syntactic processing between late L2 learners and native speakers and among learners is proficiency.

In this study, we recorded Event-Related Potentials (ERPs), which represent changes in brain electric activity triggered by particular events (Fabiani et al., 2000), a method frequently used to study language processing on-line due to its high temporal resolution. Two main components have been identified for the study of syntactic processing: the LAN (Left Anterior Negativity), a negative shift sensitive to difficulties in morphosyntactic processing; and the P600, a centro-posterior positive shift triggered by a large variety of syntactic violations and thought to reflect control and reanalysis processes (Hahne and Friederici, 2001; Tanner et al., 2013). This study explored the effect of proficiency on the native-likeness of syntactic processing in adults having learned an L2 mostly at school. Advanced learners in studies have generally been living in the country for several years; here our population corresponds to typical advanced French students, who have attained the highest level of University education in English but have spent at most an academic year in an L2-speaking country.

## Methods

### Participants

12 English native speakers (NS) as well as 12 intermediate (IS) and 12 advanced (AS) French learners of English took part in the experiment. IS and AS differed in length of University education in English (1 vs 7 semesters), time spent in English-speaking country (< 2 weeks vs 1 year) and scores at a short proficiency test.

### Material and Procedure

The material consisted of 80 short active sentences composed of the pronoun *He*, a verb and a short complement, half of which containing a violation of subject-verb agreement. 80 additional sentences with the pronoun *They* were included to balance the presence of the *-s* morpheme at the end of the verb between correct and incorrect conditions.

A fixation cross appeared first for 1000 ms and remained on the screen during the auditory presentation of the stimulus and for 1000 ms after its end. A screen prompted the participant to evaluate the grammaticality of the sentence by pressing a coloured button and remained for at most 2000 ms.

### EEG data acquisition and analysis

EEGs were recorded with a Biosemi ActiveTwo system with 32 active electrodes, referenced on-line to the two mastoids and re-referenced off-line to the average of the two mastoids. Data were filtered on-line between 0.1 and 100 Hz. Electrode impedance was maintained below 20 Ohms and the signal was sampled at a rate of 512 Hz. Epochs from -200 ms to 900 ms around the critical point (beginning of the word following the violation) were extracted from continuous data. After baseline correction (-200-0 ms), high-pass filtering at 0.1 Hz and low-pass filtering at 30 Hz, trials for which peak-to-peak amplitude exceeded 70  $\mu$ V on the EOG channel or 150  $\mu$ V on the other channels were automatically rejected. Electrodes were divided into central and lateral sites, the latter also divided into anterior/posterior region and left/right hemisphere. The following temporal windows were selected: early (500-700 ms) and late (700-900 ms) P600, and two windows for early negativities (200-300 ms and 300-500 ms).

## Results<sup>1</sup>

A P600 effect was obtained for all groups: in all the windows for NS (500-700 ms, central electrodes:  $F(1,55) = 20.02$ ,  $p < .001$ ), in the late one at central electrodes for AS ( $F(1,55) = 4.83$ ,  $p < .05$ ) and in the early one at lateral sites for IS ( $F(1,77) = 6.12$ ,  $p < .05$ ). Additionally, the Condition \* Region interaction was significant for NS at lateral sites ( $F(1,77) = 4.05$ ,  $p$

<.05), the effect of Condition being limited to the posterior region ( $p < .001$ ). Between 200 and 300 ms, NS and AS exhibited a significant negativity in the difference wave (NS, central sites:  $F(1,55) = 13.35$ ,  $p < .001$ ; AS, central sites:  $F(1,55) = 26.81$ ,  $p < .001$ ), an effect which continued for AS into the 300-500 ms window (central sites:  $F(1, 55) = 45.56$ ,  $p < .01$ ). For IS on the contrary, a positivity was observed in this window (central sites:  $F(1,55) = 8.65$ ,  $p < .01$ ). The effect of the Group \* Region interaction was significant in the 500-900 ms window at lateral sites ( $F(2, 99) = 5.41$ ,  $p < .01$ ), the effect of Group being limited to the posterior region : the P600 effect was larger for NS than AS ( $p < .001$ ) and IS ( $p < .001$ ) (see Figure 1).

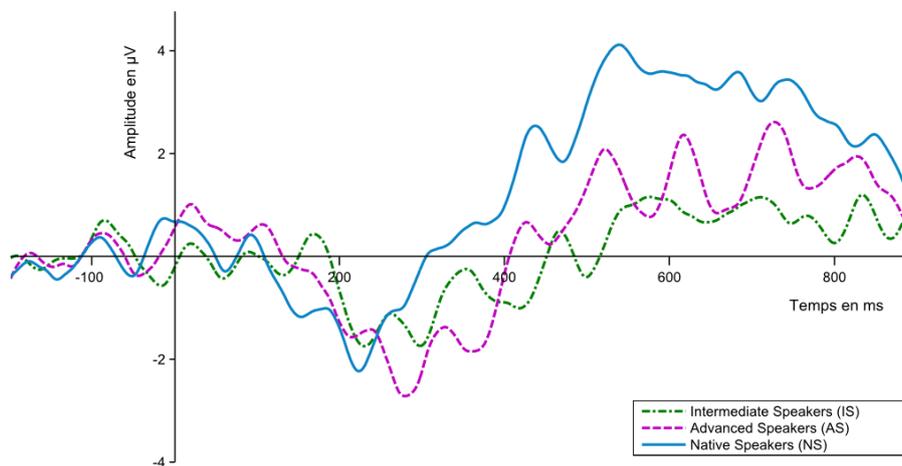


Figure 1. Difference wave (Incorrect – Correct Conditions) at PZ.

## Discussion

This study compared the ERPs evoked in native speakers and French learners of English of advanced and intermediate levels by a syntactic violation working in a similar manner in both languages, namely subject-verb agreement in the third person. A P600 effect was obtained in all groups, but with a larger amplitude for native speakers than learners, even the advanced ones. This difference reflects the difficulty learners experience in reanalysing syntactic mismatches in their L2. Analyses revealed a non-lateralized early negativity in response to agreement violations in native speakers, demonstrating an early identification of a morphosyntactic violation. This negativity was also observed in advanced learners but not in the intermediate ones.

These findings show a proficiency effect only in early ERP responses to syntactic violations but not in reanalysis processes – which have been found

to attain qualitative native-likeness as early as after six months of instruction (Osterhout et al., 2006). The fact that the advanced learners in this experiment exhibited a quantitatively non-native-like P600 questions their capacity to attain native-like processing in the L2. However, this ability has been demonstrated before for structures that work in a similar way in the learners' first and second languages, as is the case here (Ojima et al., 2005; Rossi et al., 2006). Since our advanced group had learned English mostly through instruction, our findings are consistent with the claim that a certain – higher – degree of exposure and implicit learning (Morgan-Short et al., 2012) is necessary to show native-like processing of morphosyntactic violations.

## Notes

1. Only analyses with significant results for at least one group are reported. *F* values in tables are reported with the following significance code: \*:  $p < .05$ , \*\*:  $p < .01$ , \*\*\*:  $p < .001$ . Cond. stands for the Condition variable.

## References

- Fabiani, M., Gratton, G., Coles, M.G.H., 2000. Event-related brain potentials: Methods, theory. *Handbook of psychophysiology* 53–84.
- Hahne, A., Friederici, A.D., 2001. Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition* 4, 123–141. doi:10.1017/S1366728901000232
- Morgan-Short, K., Steinhauer, K., Sanz, C., Ullman, M.T., 2012. Explicit and Implicit Second Language Training Differentially Affect the Achievement of Native-like Brain Activation Patterns. *Journal of Cognitive Neuroscience* 24, 933–947. doi:10.1162/jocn\_a\_00119
- Ojima, S., Nakata, H., Kakigi, R., 2005. An ERP study of second language learning after childhood: effects of proficiency. *Journal of Cognitive Neuroscience* 17, 1212–1228. doi:10.1162/0898929055002436
- Osterhout, L., Mclaughlin, J., Pitkänen, I., Frenck-Mestre, C., Molinaro, N., 2006. Novice Learners, Longitudinal Designs, and Event-Related Potentials: A Means for Exploring the Neurocognition of Second Language Processing. *Language Learning* 56, 199–230.
- Rossi, S., Gugler, M.F., Friederici, A.D., Hahne, A., 2006. The Impact of Proficiency on Syntactic Second-language Processing of German and Italian: Evidence from Event-related Potentials. *Journal of Cognitive Neuroscience* 18, 2030–2048. doi:10.1162/jocn.2006.18.12.2030
- Tanner, D., Mclaughlin, J., Herschensohn, J., Osterhout, L., 2013. Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition* 16, 367–382. doi:10.1017/S1366728912000302.

# Perception and production of Italian L2 sounds

Luciano Romito, Manuela Frontera

Department of Languages and Education, University of Calabria, Italia

## Abstract

Various theories and models have been argued and experimented in order to explain both the phonological and phonetic influence or interference of native language on perception and interpretation of L2 categories (Vayra *et al.* 2012; Kuhl *et al.* 2006), and whether or how much production depends on it (Bohn & Flege 1990). The oppositions examined here deal with vowels (final unstressed /e – i/, /o – u/, an opposition that in Galician language is reduced to /e – o/) and consonants (/b – v/, /ts – dz/, both absent in Galician). The goals are: (a) to verify levels of identification of Italian non-native vowel and consonant oppositions by Galician students of Italian L2; (b) to analyse discrepancies in production between Italian native speakers and Italian L2 learners.

Key words: applied linguistics, L2 learning, phonetic perception, production, PAM

## Identification experiments

### Method

#### Participants

Twenty-four undergraduate students of the University of Santiago de Compostela (Galicia, Spain), without history of severe language or hearing impairments, participated in this study. They were all students of Italian L2, twelve since one/two year/s, twelve since three/four. All students were Galician but some of them use Castellan as L1.

#### Stimuli and procedures

Two identification tests were proposed to the students using Folerpa (Fernández Rei *et al.* 2014). Stimuli were recorded by three native Italian speakers (2F, 1M from northern and central Italian varieties). They consisted of words pronounced in the phrasal context “dico X rapidamente/con calma” (“I say X rapidly/with calm”) where X is the target word. Both real and nonsense words were used. The target words were then isolated and normalized by an editing software. Each test consisted of 180 tokens (9 minimal pairs, 4 oppositions, 5 repetitions per word) divided into four blocks. The first test dealt with vowel oppositions of final unstressed /e – i/, /o – u/, the second with consonants /b – v/, /ts – dz/. Every token was repeated twice after 1 sec. interval and presented with no feedback. Time responses were also evaluated. Results to the tests were compared with those of a control group of six native Italians from the University of Calabria.

## Results and discussions

### Vowels

The relations between correct identifications and years of Italian L2 experience or first language of subjects (Galician/Castellan) were not significant according to chi-square tests ( $p = 0,036$ ;  $0,0001$ ) and one-way ANOVA results. Subjects successfully identified final unstressed /e/ and /o/ (native oppositions), better than /i/ and /u/ (non-native oppositions), as well as the control group did: this could reasonably depend on an intrinsic poor quality realization of unstressed final vowels by the Italian speakers (Romito *et al.* 1997). Time responses means confirm the tendency to a lower identification of high vowels.

### Consonants

Statistical analysis reveal slightly significant differences between Galician and Castellan L1 speakers: the lasts seem to identify better all consonant oppositions, even if both groups obtain percentages of success worse than the previous test (79-98% to 57-90%). Percentages decrease when associated to non-native sounds /v/, /ts/ and /dz/, where the affricate /ts/ gets the worst results, while the native sound /b/ is almost often correctly identified. Time responses show troubles in the identification of voiced/unvoiced affricates, as well as for the Italian control group: this may confirm an unaware distinction in the use of the two phonemes by Italian speakers, due to the unpredictability of this specific opposition (Gili Fivela 2010).

## Production

### Method

#### Participants

Seven female Italian L2 students of the University of Santiago de Compostela and two Italian native speakers. All students were Galician and use Galician as a first language. Native speakers productions were the same used for perceptual experiments.

#### Corpus

All speakers produced 30 phrases per vowel opposition, containing 15 minimal pairs of target words (ex. ‘Dico *molti* rapidamente, dico *molte* rapidamente’; ‘dico *pirco* con calma, dico *pircu* con calma’). Target words were isolated by an editing software. Final unstressed vowels were labeled and analysed by Praat (5.4.09). Values of F<sub>1</sub>, F<sub>2</sub> and duration were extracted by means of a script adapted by Luciano Romito.

### Results and discussions

Italian L1 productions of final unstressed vowels reflect processes of intrinsic centralization in front vowels areas and neutralization in back ones:

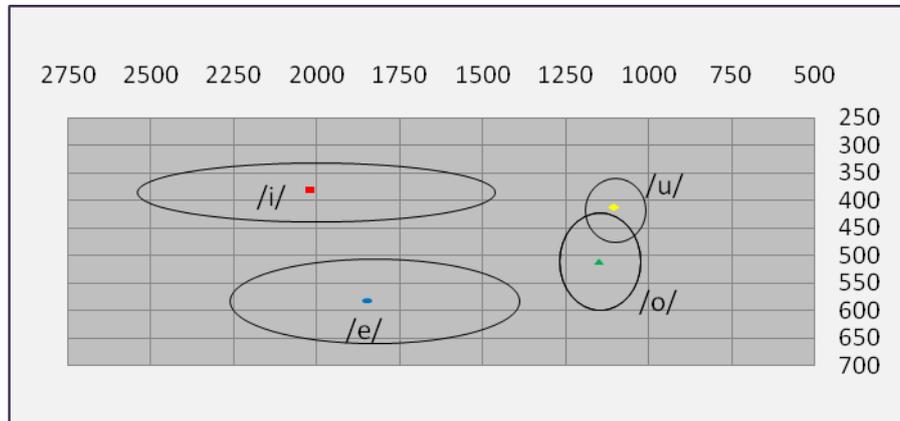


Figure 1: Italian unstressed vowels areas.  $F_2$  values are on the x-axis,  $F_1$  values are on the ordinate.

Galician speakers ‘polarize’ front vowels /i – e/ and back /u/, while close-mid /o/ is almost centralized: these results show both likely efforts in producing the opposition between close and close-mid front vowels, and a strong influence of Galician L1 close-mid vowels production, almost evident in /o/, extremely centralized (Regueira 2003). Extra-fronted realization of /i/ and /e/ reflects longer durations of Italian L2 vowels respect to the Italian L1 ones. On the contrary, Italian L2 back vowels duration is slightly reduced.

Ita_L1				Ita_L2			
	F1	F2	(ms.)		F1	F2	(ms.)
[e]	586	1927	61.783	[e]	538	2074	69.882
[i]	389	2003	62.402	[i]	404	2425	75.166
[o]	511	1147	52.729	[o]	459	1045	45.268
[u]	422	1072	46.157	[u]	415	828	43.735

Table 1: Values of  $F_1$ ,  $F_2$  and duration of Italian L1 and L2 final unstressed vowels produced by female subjects.

## Conclusion

The perception of sounds belonging to an L2 is partly influenced by the phonetic and phonological features of the native language. This study reveals how the perception of native sounds is almost always correct compared to that of nonnative ones. However, the proximity of the two linguistic systems usually guarantees high percentage scores of success. As far as production is concerned, the current work is at the moment focused on vowels, but it will soon include an analysis on the consonant oppositions perceptually tested. By now, the production of nonnative vowel oppositions seems to confirm an ability in discrimination, tested by the perceptual experiments, although it appears rather unnatural. In this sense, it is possible to confirm that L1 not only affects perception but also production in L2: as a consequence, learners will be able to produce L2 oppositions, but with acoustic values almost near to the equivalent native sounds. Further analysis on the connection existing between perception and production will be led on consonants, which seemed to be more problematic for our subjects.

## References

- Bohn, O.-S., Flege, J. 1990, Interlingual identification and the role of foreign language experience in L2 vowel perception in *Applied Psycholinguistics*, 11, 303-328.
- Fernández Rei, E. (coord.), 2014, FOLERPA: Ferramenta On-Line para ExpeRimentación PerceptivA. Santiago de Compostela: Instituto da Lingua Galega. <http://ilg.usc.es/FOLERPA/>
- Gili Fivela, B. 2010, Affricate, in *Enciclopedia dell'Italiano*, R. Simone (dir.), Istituto della Enciclopedia Italiana G. Treccani, Roma.
- Regueira, X. L. 2003, Vocais finais en galego e en portugués: un estudio acústico, *Actas do VII Congreso Internacional de Estudos Galegos* (Barcelona, 28-31 de maio de 2003).
- Romito L., Turano T., Loporcaro M., Mendicino A. 1997, Micro e Macrofenomeni di centralizzazione nella variazione diafasica: rilevanza dei dati fonetico-acustici per il quadro dialettologico del calabrese in *Atti del convegno VII Giornate di Studio del Gruppo di Fonetica Sperimentale* (Napoli, novembre 1996).
- Tsao, F. M., Liu, H. M., Kuhl, P.K. 2006, Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants, *The Journal of the Acoustical Society of America* 120 (4), 2285-2294.
- Vayra, M., Avesani, C., Best, C. T., Bohn, O. S. 2012, Non solo dettagli fonetici, non solo categorie fonologiche: l'interazione tra fonetica e fonologia nella percezione di suoni non-nativi, in *Studi e Saggi Linguistici*, Pisa, ETS.

# Audiovisual processing of pharyngealization and length in Emirati Arabic

Mark Scott<sup>1</sup>, Ali Idrissi<sup>2</sup>

<sup>1</sup>United Arab Emirates University, Al Ain, Abu Dhabi, UAE

<sup>2</sup>Qatar University, Doha, Qatar

## Abstract

This paper addresses a gap in the literature on audiovisual speech perception. Previous research has examined how the perceived primary place-of-articulation of a speech sound is influenced by visual information. Visual influences on the perceived length of a speech sound or on the presence of pharyngealization have not been examined. The experiments reported here demonstrate that the perception of both singleton/geminate and pharyngealized/plain is susceptible to visual influence.

Key words: Audiovisual, perception, geminate, pharyngealization, Arabic

## Introduction

Most studies of the McGurk effect have demonstrated that the perceived place-of-articulation of a sound can be shifted by visual information. Arabic has contrasts that are not so clearly about (primary) place-of-articulation. Geminates involve holding a constriction for a longer time; pharyngealization involves a secondary constriction which is not itself visible. The two experiments below test whether Arabic speakers are visually sensitive to these contrasts. The prediction in **Experiment One** is that the consonant length in a video will shift perceived *auditory* consonant length to match. In **Experiment Two**, we predict that video of pharyngealized or plain sounds will shift *auditory* perception of pharyngealization to match –here the perceiver is not seeing the pharyngeal gesture, but its visual effects on surrounding sounds.

## Experiment One: Audiovisual Perception of Geminaton

In comparison to hearing, vision is poor at timing perception. This means that when there is a conflict, we generally perceive timing in line with information from hearing (Goldstone & Lhamon, 1974). This is the modality appropriateness hypothesis (Welch, 1999) and it suggests (contra this experiment) that visual timing influences on hearing are unexpected.

## Procedures

An audio continuum between /abah/ (singleton) and /abbah/ (geminate) was created by morphing between clear endpoints (recordings of a native Emirati

Arabic speaker). The same speaker was video recorded and 8 videos each for singleton and geminate (free of any obvious irregularities) were chosen for use. For both singleton and geminate conditions, two interleaved audio staircases were presented with random switching. Each staircase had 12 turns. For the first 6 turns of each staircase, the audio was presented alone; after 6 turns, the audio was accompanied by videos of either singleton or geminate. There were 40 native Arabic speaking female participants.

### Results & Discussion

Each participant's data was fit to a psychophysical curve.

participants' data were unfittable and were eliminated from analysis. A paired t-test comparing the fitted thresholds between "Long" and "Short" was significant [ $t = -3.1105$ ,  $df = 32$ ,  $p = 0.0039$ ]. This demonstrates that people can integrate audiovisual speech timing information. This is surprising, as the auditory system typically dominates the audiovisual perception of timing.

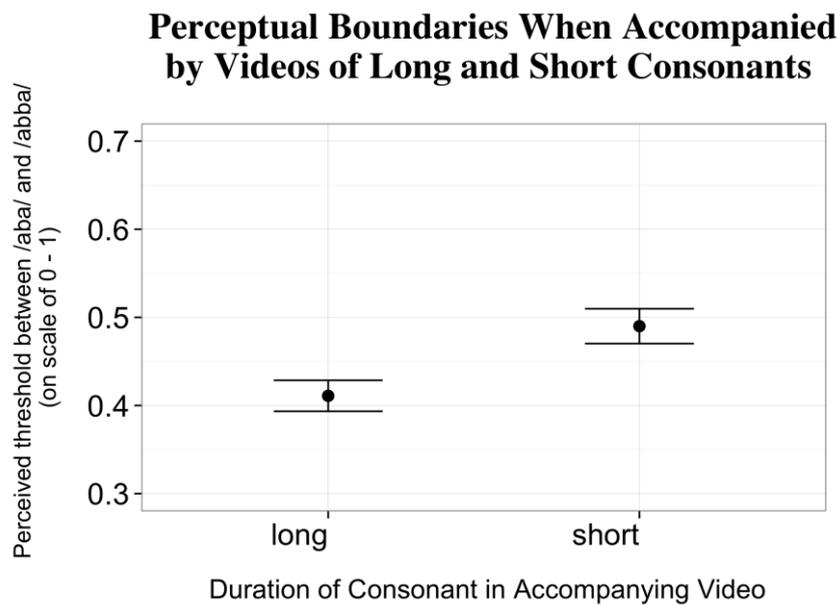


Figure 1. Results of Experiment One – Standard-Error bars are shown.

## Experiment Two: Pharyngealization

Most Arabic dialects have *pharyngealized* consonants. This is a secondary articulation that involves a constriction at the upper part of the pharynx (Ladefoged & Johnson, 2011). Pharyngealization typically spreads to neighbouring segments (Zawaydah, 1999). In Gulf Arabic, this pharyngealization is accompanied by some degree of lip rounding (Watson, 1999).

### Procedures

This experiment was identical in structure to **Experiment One** except that the sounds were /aḏah/ (plain) and /aḏʰah/ (pharyngealized). The audio continua and videos were created in the same way as in **Experiment One**. The same participants as in **Experiment One** were run in this experiment.

### Results & Discussion

Each participant's data was fit to a psychophysical curve. 15 participants' data were unfittable and were eliminated from analysis. A paired t-test comparing the fitted thresholds between "Pharyngealized" and "Plain" was significant [ $t = -2.4992$ ,  $df = 24$ ,  $p = 0.0197$ ]. As predicted, more pharyngeal were heard with a pharyngeal video and similarly for plain.

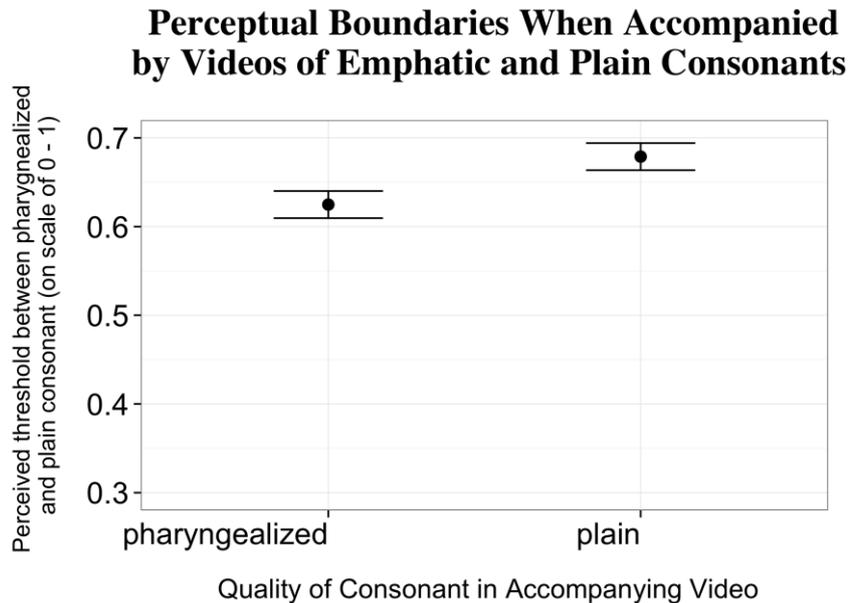


Figure 2. Results of Experiment Two – Standard-Error bars are shown.

This shows that Arabic speakers are sensitive to visual cues to pharyngealization, and can integrate these cues into a combined audiovisual percept. Clearly, participants are not seeing the secondary constriction at the pharynx itself, as this is necessarily occluded. Instead, pharyngealization is being detected via other visual cues such as jaw-retraction and lip rounding.

## Conclusions

The use of visual cues to length and secondary articulations has not received much attention. These experiments address this gap by showing that visual information about the duration (singleton versus geminate) of a consonant and the presence of a secondary articulation (pharyngealization) on segments can be detected visually and integrated into an audiovisual percept.

## Acknowledgements

This research was supported by UAEU Grant 31h076. We would like to thank our Assistants: Souad Al Helou and Hamdha Saleh Salem Alkarbi.

## References

- Goldstone, S. and Lhamon, W.T. 1974. Studies of auditory-visual differences in human timing judgment: Sounds are judged longer than lights. *Perceptual and Motor Skills* (39), 63–82.
- Ladefoged, P. and Johnson, K. 2011. *A Course in Phonetics* 6th ed., Cengage Learning, Wadsworth.
- Watson, J.C.E. 1999. The Directionality of Emphasis Spread in Arabic. *Linguistic Inquiry* 30(2), 289–300.
- Welch, R.B. 1999. Chapter 15 Meaning, attention, and the ‘unity assumption’ in the intersensory bias of spatial and temporal perceptions. In Aschersleben, G., Müssele, J. & Bachmann, T. (eds.), *Advances in Psychology*, 371–387, New York, North-Holland.
- Zawaydah, B.A. 1999. *The phonetics and phonology of gutturals in Arabic*. Ph.D. Indiana: Indiana University.

# Some effects of Vowel Space Area (VSA) reduction in speech intelligibility

Patrizia Sorianello

Department of Letters, Languages and Arts. University of Bari, Italy

## Abstract

This study intends to verify the presence of a correlation between the size of vowel space and the speech intelligibility. A reduced intelligibility is a common problem for those individuals suffering from dysarthria. They show a compressed vowel space with an important degree of acoustic centralization. A recent metric employed to distinguish healthy from dysarthric speech is represented by the Vowel Space Area (VSA). We have tested this parameter by comparing subjects with Down Syndrome to control speakers. The results of the present research, the first available data for Italian language, proved that the VSA is a suitable acoustic predictor of an impaired intelligibility.

Key words: speech intelligibility, Vowel Space Area, Down Syndrome, dysarthria

## Introduction

Intelligibility refers to the degree of comprehensibility of speech. A reduced intelligibility is a typical aspect of voice and speech disorders, irrespective of the underlying neurological condition. Most types of dysarthria are characterized by the loss of vowel clarity, altered range of articulatory movements and vowel centralization. The methodological procedures used to estimate the intelligibility include parental surveys, tests based on the auditive identification of words and scoring from transcriptions. Previous studies have shown a positive relationship between the size of vowel system and the overall intelligibility (Weismer et al 2001, Liu et al. 2005). Larger vowel space are expected to be associated in more intelligible subjects, while smaller vocalic system are expected for dysarthric ones. Vowel Space Area (VSA) is an acoustic metric able to distinguish dysarthric patients from healthy control ones. VSA is generally constructed by the F1/F2 values of the three corner vowels /i/, /a/ and /u/. It is calculated by means of the equation reported in (1):

$$(1) \quad \text{triangular VSA: } [(F1i (F2a-F2u) + F1a (F2u-F2i) + F1u (F2i-F2a)]/2$$

Nevertheless, the predictive force of VSA is discussed, since sometimes it failed. The VSA seems to depend on several external factors such as the degree of inter-speaker variability or the number of the vowels considered. For this reason, in recent studies other metrics have been also considered (for a review, Lansford & Liss, 2014).

## Speech intelligibility in Down Syndrome

In this research we focus our attention on individuals with Down Syndrome (DS). They exhibit an atypical language with a degraded intelligibility. Generally, the speech of people with (DS) is difficult to understand: phonetic alterations, dysfluency (stuttering or cluttering), altered voice quality and dysprosody are the most frequent manifestations of an impaired intelligibility (among all, Nadel 1998, and for Italian language Sorianello 2012). In DS the precision of the articulatory gestures involved in the phonation is altered by several structural restrictions such as facial hypoplasia, hearing difficulties, breathing dysfunctions, memory deficits. In addition, speech intelligibility is also affected by impairment in both oral motor control and planning (Kumin et al. 2006). With regard to the vowels, the presence of a smaller vowel space has been observed in DS as compared to matched control group (Moura et al. 2008, Bunton & Leddy 2011). Speech intelligibility of Italian people with DS has never been instrumentally analyzed. The aim of the present study was to examine the effects of a reduced speech intelligibility on vowel working space.

## Methodology

Two groups of Italian subjects took part in this study. Ten young adults with DS (5 M, 5 F), aged between 16 and 22 and lived in Cosenza, a city of Southern Italy, were compared to age-matched controls (3 M, 3 F) (CG). A corpus of 1200 stressed vowels was analyzed. It consisted of *Isolated Words*; subjects were asked to name 108 pictures presented on cards representing common objects or actions. All participants were audio recorded using a digital TASCAM DR-07; for people with DS we previously obtained informant consent from the family. For each vowel the following parameters were extracted by means of the software PRAAT: 1) F1, F2 frequency values, 2) Centroid value, 3) triangular VSA, 4) coefficient of variation (CV: Standard deviation/mean). The formants were manually extracted using FFT procedure, at the correspondence of the steady-state portion of the formant trajectory. A one-way Anova test was performed.

## Results

In Figures 1-2 the VSA of DS and CG subjects are graphically shown; the mean values of the VSA are listed in Table 1. In DS the vowel space is relatively small, above all in males. As expected (see Table 1), the mean values of the VSA are lower in males than females, in both groups. Compared to control speakers, the VSA size of subjects with DS is always smaller; a significant group effect (DS vs. CG) is found ( $p = .001$ ). In detail, in males the size of VSA is 61% smaller with respect to the CG, whereas in

females the difference amounts to 44,5% (males vs. females  $p=.001$ ). Data are consistent with a condition of altered speech intelligibility, suggesting that the VSA is a valid metric to capture differences between groups.

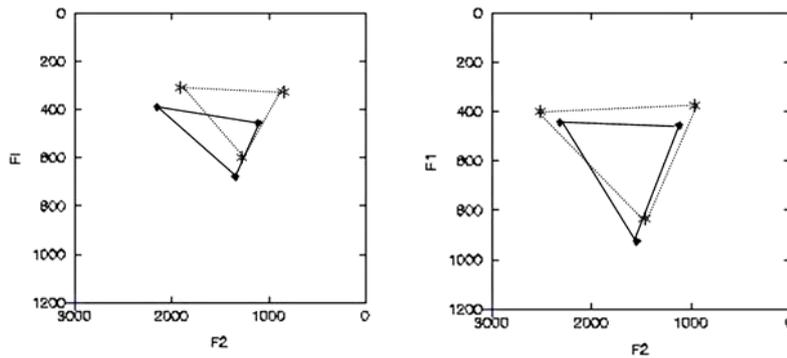


Figure 1. VSA of individuals with DS (closed line) and CG individuals (dashed line), grouped into males (left) and females (right)

Table 1. Mean, standard deviation (SD) and coefficient of variation (CV) of the triangular VSA for DS and CG. VSA and SD are expressed in  $\text{Hz}^2$ .

Vowel Space Area				
Values	DS		CG	
	males	females	males	females
Mean	125.144	240.161	201.508	347.052
SD	12.171	55.779	54.452	38.446
CV	9.7%	23.2%	27%	11,1%

The coefficient of variation is less informative, being lower in males with DS, but relatively large in females. The degree of vowel dispersion is likely due to a certain inter-speaker variability. Overall, in DS the dispersion of front vowels is normally larger than that of back ones, due to the difficulties encountered by speakers with DS in establishing accurate lingual contacts.

## Conclusions

Experimental results demonstrate that speakers with DS are effectively distinguished from healthy ones by means of the triangular VSA. There is a strong relationship between VSA and vowel distinctiveness: the smaller the area, the less intelligible the speech. So, at least in our data, the VSA is able of differentiating people with or without dysarthria and can be taken as a quantifier of vocal abnormality. Contrary to what was observed in other studies, an impaired intelligibility is anchored to the presence of a restricted VSA. This result can probably be attributed to the fact that all participants

with DS show a significant level of dysarthria. People with DS meet hard difficulties in performing speech sounds. Their anatomical anomalies have strong implications on language production, subjects with DS have a more reduced F1/F2 plane than CG individuals. The presence of a large tongue, in relation to a smaller oral cavity, produces inadequate movements of tongue, lips and jaw. The inability to synchronise oral gestures is the cause of frequent processes of phonological target undershoot. Moreover, the weakness of the oro-facial muscles, a lack of laryngeal control and motor disorder inevitably diminishes the accuracy of the speech sounds. Given this starting situation, it is not surprising that in DS the precision in achieving a specific articulatory target is strongly reduced. Their vocalic system shows a widespread dispersion: all the vowels share common areas. Very often, the vowels are too close within the system, so their articulatory identity becomes unclear and that interferes with the possibility to establish effective perceptive contrasts. From a phonetic point of view, a great deal of stressed vowels are indistinct sounds. This situation induces a consistent compression of the vowel space with clear vowel alterations. As a consequence, when the articulation is damaged, speech intelligibility automatically decays. In order to confirm our findings, in the future we intend to perform a more complete assessment of vowel articulation, exploring the impaired intelligibility by means of other metrics, such as the Formant Centralization Ratio (FCR) or the Vowel Articulation Index (VAI).

## References

- Bunton, K. and Leddy, M. 2011. An evaluation of articulatory working speech area in vowel production of adults with Down syndrome. *Clinical Linguistics and Phonetics* 25, 321-334.
- Kumin, L. 2006. Speech intelligibility and childhood verbal apraxia in children with Down Syndrome. *Down Syndrome Research and Practice* 10, 10-22.
- Lansford, K.L. and Liss, J.M. 2014. Vowel acoustics in dysarthria: speech disorder diagnosis and classification. *Journal of Speech, Language and Hearing Research* 57, 57-67.
- Liu, H.M., Tsao, F.M. and Kuhl, P.K. 2005. The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *Journal of Acoustical Society of America* 117, 3879-3889.
- Moura, C.P., Cunha, L., Vilarinho, H., Cunha, M., Freitas, et al., M. 2008. Voice parameters in children with Down Syndrome. *Journal of Voice* 22, 34-42.
- Nadel, L. (ed.). 1988. *The psychobiology of Down syndrome: issues in the biology of language and cognition*, Cambridge (Mass.), MIT Press.
- Sorianello, P. 2012. *Linguaggio e Sindrome di Down*, Milano, FrancoAngeli.
- Weismer, G., Jeng, J.Y., Laures, J., Kent, R. and Kent, J. 2001. Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatica et Logopaedica* 53, 1-18.

# The intonation of Albanian polar questions and statements

Charalambos Themistocleous<sup>1</sup>, Daniela Müller<sup>2</sup>

<sup>1</sup>Department of English Studies, University of Cyprus

<sup>2</sup>Institut für Phonetik und Sprachverarbeitung, Ludwig-Maximilians-Universität München

## Abstract

This study aims to provide an account of the effects of sentence type (statements vs. polar questions) on Standard Albanian prenuclear rises through a polynomial model representing the dynamic characteristics of tonal contours. Results show that the main difference in contour shape between Albanian statements and polar questions is located in the shape of the prenuclear rise, and this difference was significant; onset timing of the prenuclear rise, however, did not differ significantly between the two types of sentence.

Keywords: Albanian, intonation, statements, questions.

## Introduction

The tonal curve consists of tonal events such as peaks and troughs that are perceived as pitch rises and falls. These tonal events convey informational meaning (e.g., focus and topic) or manifest sentence types (e.g., statements, questions, etc.) (Botinis, Granström & Bernd, 2001). There has been considerable research on Albanian segments, yet studies on Albanian intonation are sparse (see, however, Manzini & Savoia, 2007; Memushaj, 2009). By employing a polynomial model that represents the dynamic characteristics of Albanian tonal contours, this study is to our knowledge the first account of the effects of sentence type (statements and polar questions) on prenuclear rises in Standard Albanian.

## Method

Data for this study were gathered in March 2015 by the second author in Tirana, the capital of Albania. Ten Standard Albanian speakers (8 male, 2 female) in their mid-twenties born and raised in the broader Tirana area participated in the study. The speech materials consist of the phrase *Melina nominon Milonin* ('Melina nominates Milon') uttered as a statement and as a polar question. Question-answer pairs between the speaker and one of the two Albanian student assistants facilitated the elicitation of narrow focus. Speakers wore a head-mounted microphone connected to a H4n Handy Recorder (Zoom Corporation, Tokyo, Japan) and were recorded in a quiet room at a 44100 Hz sampling frequency. All tokens obtained from the eight

male speakers were used for the acoustic analysis (data from the female participants were obtained for a separate study). The first peak of each utterance was selected from a local minimum up to a following local minimum and analysed.

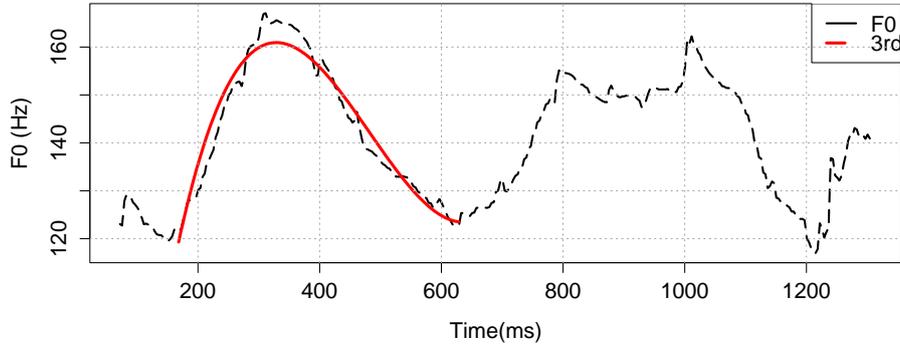


Figure 1. Tonal contour of the utterance *Melina nominon Milonin* uttered with narrow focus on the rightmost constituent by a male speaker of Albanian. The pre-nuclear peak is fitted with a 3<sup>rd</sup> degree polynomial (red line).

To provide the best fit to time-frequency samples from each tonal contour, a polynomial curve fitting method was used (see Figure 1). The resulting  $N^{\text{th}}$  degree polynomial is given by:

$$(1) \quad f_{(m)}(t) = a_{0,m} + a_{1,m}t + a_{2,m}t^2 + \dots + a_{N,m}t^N$$

where  $t$  is time. The index  $m=1,\dots,M$  represents the tonal contours approximated with curve fitting. Moreover, the distance between the beginning of the rise, i.e. the first low and the onset of the stressed syllable was measured. Note that the timing of the peak and the end of the rise are determined by the polynomial model. A  $t$ -test was conducted on the polynomial coefficients. Sentence type (statement vs. question) was entered into the model as an independent variable; we fitted a 3<sup>rd</sup> degree polynomial. The statistical analysis was carried out in R 3.0.2 (R Core Team, 2012).

## Results

*i. The overall contour.* In both statements and questions, an  $f_0$  rise-fall associates with the stressed syllable. The rise begins at a local minimum at the left edge of the stressed onset consonant and the peak aligns inside the post-stressed vowel (see Figure 2). In statements, the nuclear pitch accent associates with a peak, whereas in polar questions, the nuclear pitch accent associates with a local minimum of the  $f_0$ . The polar question contour ends in an  $f_0$  rise.

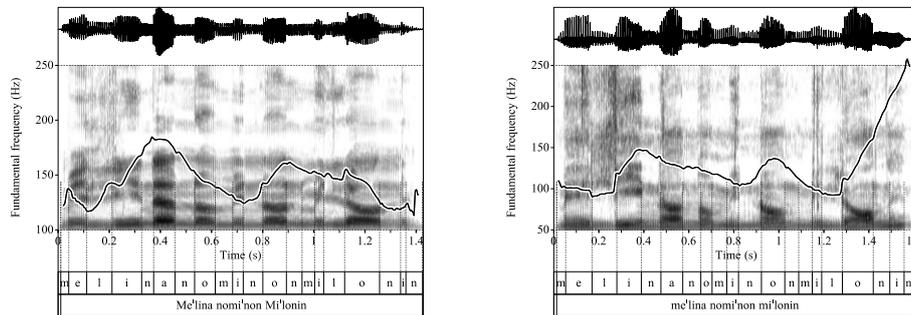


Figure 2. Tonal contour of the utterance *Melina nominon Milonin* uttered as a statement (left) and as a polar question (right) by a male speaker of Standard Albanian.

*ii. Pre-nuclear peaks.* Figure 3 shows the output of the curve fitting for the male participants. Overall, a *t*-test showed that questions vs. statements (*sentence type*) do not differ in their starting frequency estimated by the zero<sup>th</sup> coefficient ( $a_0$ ); there were also non-significant effects for the  $a_2$  and  $a_3$  coefficients. *Sentence type* had, however, significant effects on  $a_1$  ( $t(1)=5.814, p < .05$ ).

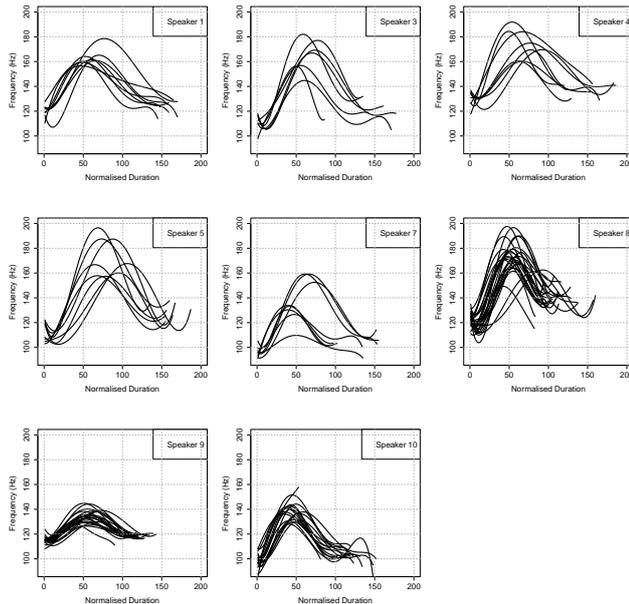


Figure 3. Fitted lines of the first prenuclear peaks for each male speaker.

*iii. Timing of pre-nuclear peaks.* In both statements and polar questions, an  $f_0$  rise-fall begins at a local minimum at the left edge of the onset consonant of the stressed syllable and the peak aligns inside the post-stressed vowel. There was no significant difference between polar questions and statements with respect to the beginning of the prenuclear rise.

## Discussion

This study used a polynomial model representing the dynamic characteristics of Albanian tonal contours in order to account for the effects of sentence type (statement vs. polar question) on Standard Albanian prenuclear rises. The results show that statements and questions differ significantly in the first coefficient, which determines the prenuclear rise. By contrast, the starting frequency of questions and statements was not significantly different. Overall, the local minimum of the prenuclear pitch accent is timed with respect to the syllable onset. The timing of the peak and the timing of the end of the rise are determined by the polynomial model. Therefore, by requiring only one anchoring point for the pitch accent, this model provides a more parsimonious account of pitch accent timing than models which require specifications for both the local minimum and the peak (see Ladd, 2008 for a review). As far as we are aware, this study provides the first model of prenuclear accents in Standard Albanian statements and polar questions.

## Acknowledgments

We thank our student assistants Amarildo Çorape and Erdit Mimani for running the experiment in Tirana and Carolin Sabath for segmenting the sound files. Moreover, we are grateful to Uwe Reichel for helpful comments on an earlier version of this paper.

## References

- Botinis, A., Granström, B., & Bernd, M. 2001. Developments and paradigms in intonation research. *Speech Communication*, 33, 263-296.
- Ladd, D. R. 2008. *Intonational phonology* (2<sup>nd</sup> ed.). Cambridge, UK, New York: Cambridge University Press.
- Manzini, R. M., & Savoia, L. M. 2007. *A Unification of Morphology and Syntax. Investigations into Romance and Albanian dialects*. New York: Routledge.
- Memushaj, R. 2009. *Fonetika e shqipes standarde*. Tiranë: Botimet Toena.
- R Core Team 2012. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>.

# The influence of tempo and speaking style on timing patterns in Polish

Agnieszka Wagner, Katarzyna Klessa

Institute of Linguistics, Adam Mickiewicz University in Poznan, Poland

## Abstract

The study investigates the effect of different speech rates and speaking styles on timing patterns in Polish using a recently developed methodology (Annotation Pro+TGA) and a new speech corpus for rhythm research. We report on preliminary results indicating changes in syllable rates according to intended tempi (i.e. higher/lower number of syllables per second at faster/slower rates) and discuss results of regression analysis of duration differences in inter-pausal time groups, e.g. the variability of duration difference regression slopes in the three speaking styles.

Key words: timing patterns, tempo, speaking style

## Introduction

The relation between timing patterns, speech rate and speech style has been the subject of extensive research for many languages, but not for Polish – the exceptions are the recent studies by Malisz (2013) who investigated tempo-differentiated utterances using coupled oscillator model or by Klessa & Gibbon (2014) and Yu et al. (2014) who looked at acceleration and deceleration patterns within interpausal groups of syllables. Malisz observed the general dominance of the syllable oscillator (pointing to syllable isochrony), especially with increasing speech rate, and a tendency towards greater stress-timing at slower rates. The findings concerning acceleration and deceleration patterns showed significant differences in duration regression slopes and intercepts between read and conversational speech styles.

The goal of the present study is to investigate the effects of speech rate and style variation on timing properties of Polish and to investigate the relationship between intended and laboratory measured tempo. For this purpose we apply a recently developed methodology of comprehensive, multidimensional speech timing analysis – Time Group Analysis (TGA) proposed by Gibbon (2013) and its implementation as Annotation Pro+TGA (Klessa & Gibbon, 2014) to a new speech corpus created specifically for the needs of speech rhythm research and consisting of phonotactically, prosodically and rhythmically differentiated speech material (isolated sentences, text “The North Wind and the Sun”, excerpts of poems, spontaneous speech and mini-dialogues) produced by 20 native and 15 non-native speakers of Polish.

## Speech material & methods

The present analyses were based on recordings of 5 native speakers of Polish reading a text (*text* subcorpus), sentences consisting of 17-22 syllables (*sent* subcorpus) and poetic verses representing different metres (*poe* subcorpus) in five different intended tempi (very slow to very fast, referred to as: *vslow*, *slow*, *norm*, *fast*, *vfast* further in the paper).

In order to investigate the possible differences in timing variability in the three types of texts and at the five different tempi, we applied the TGA approach which is based on computational annotation mining procedures and provides information on the perceptually significant properties of the temporal structure of utterances including isochrony, iteration and alternation. The main TGA parameters are the *slope* (representing speech acceleration and deceleration) and *intercept* regression values.

The speech material was annotated with Annotation Pro (Klessa et al., 2013). First, the recordings were semi-automatically labelled on the orthography level, then a GTP conversion was performed on the level of phones, syllables and words using Salian plugin for Annotation Pro. Afterwards, all segment boundaries were inspected by trained labellers, and manually adjusted in case of a need. Finally, the data were automatically segmented into inter-pausal time groups for the needs of TGA, using the Annotation Pro+TGA plugin (Klessa & Gibbon 2014).

## Results

Altogether, 1265 inter-pausal groups were distinguished in the present material. The number of the groups varied among the speakers (min. average of 38 groups, max. average of 64 per speaker) and among the different tempi with more groups produced at slower than at faster rates.

The mean values of duration difference *slopes* obtained for the recordings at the five tempi for the three different speaking styles are shown in Figure 1. Generally, the slopes tended to be slightly higher at normal or slower rates than at the very fast or fast rates, depending on the speaking style. The low values of slope in the *vfast* condition (and also in the *norm* condition for the *sent* subcorpus) indicate that speakers realized this part of the recording task not only by speeding up on average, but also by applying less deceleration within the time groups. More deceleration, i.e. more dynamics in within-group timing can be found at slower rates. When comparing the speaking styles, greater differences in slopes are present in the utterances from the *poe* and *text* subcorpora than in the isolated sentences – this might be attributed to the wider range of possibilities given by longer texts or poetic verses, e.g. in terms of interpretation by individual speakers.

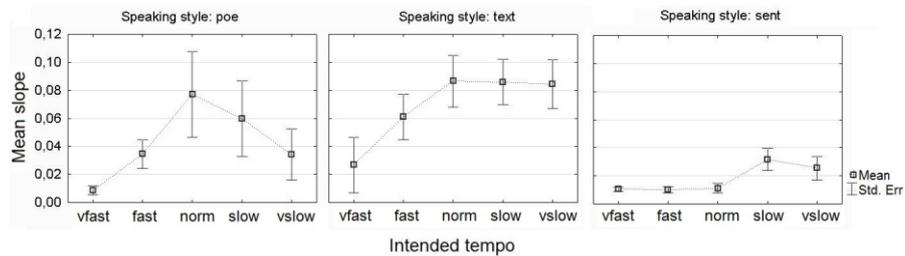


Figure 1. Mean values of duration difference slopes.

The speakers achieved the targets assumed for the five intended tempi by producing utterances significantly differing in terms of rates in these five conditions for each of the three speaking styles. The mean speaking rates (measured in syllables per second) realized by the speakers were in agreement with the assumptions of the recording scenarios, and were the fastest in the *vfast* condition and the slowest in the *vslow* condition. Interestingly, when the speaking rates were normalized by speakers and files, it was observed that on average, the rates were slightly higher in the *sent* subcorpus than in poetic verses and read text (Figure 2).

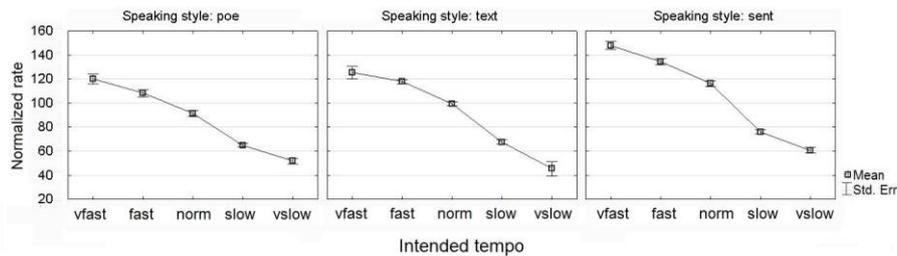


Figure 2. Mean values of speech rates (syll/sec) realized by the speakers.

The mean values of regression *intercepts* are expected to correlate with segmental durations and thus – to negatively correlate with speaking rates. This was also generally confirmed for the present data (Figure 3).

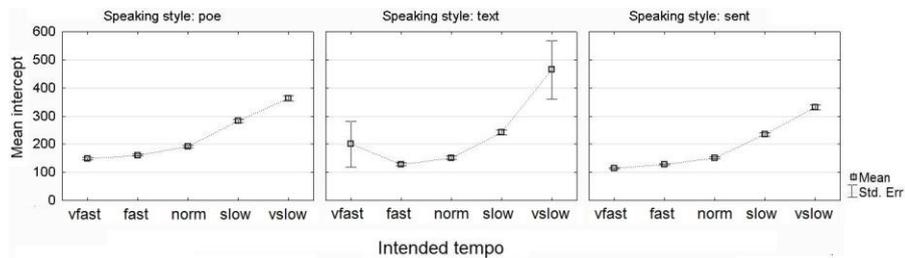


Figure 3. Mean values of duration difference intercepts.

An exception is the *text* subcorpus where the mean intercept in the very fast tempo was relatively higher than in the *fast* and *norm* conditions. However, it should also be noted that the result in the *vfast* condition had the highest standard error (and standard deviation) values, which shows more individual differentiation in this speaking style. The largest dispersion around the mean was found for the *vslow* condition, also in the *text* subcorpus.

## Discussion and conclusions

The present paper provides the first insights into a new corpus designed primarily for analyses of durational and rhythmic phenomena in Polish. As it has been discussed, the variability of the actual average values of speech rates was generally in agreement with the assumed five tempo-related conditions. The intercept values were negatively correlated with speaking rates and the patterns were similar across speaking styles (except for *vfast* condition in the *text* subcorpus). The highest variability was found in mean slope values, indicating that the deceleration and acceleration patterns vary more among the speaking styles than intercepts or mean speech rates.

Due to the limited amount of speech material used, the present paper should be treated as a pilot study. Further work will include a replication of the present analyses based on the whole corpus, as well as a series of follow-up timing analyses with the use of other methods and rhythm metrics.

## Acknowledgements

This work was supported by National Science Centre (NCN) grant no. 2013/11/D/HS2/04486.

## References

- Gibbon, D. 2013. TGA: a web tool for Time Group Analysis. Proc. of TRASP. In Hirst, D. & Bigi, B. (eds.), Proc. of TRASP, 66-69, Aix-en-Provence, France.
- Klessa, K. and Gibbon, D. 2014. Annotation Pro + TGA: Automation of speech timing analysis. Proc. of LREC, Reykjavik, Iceland.
- Klessa, K., Karpiński, M. and Wagner, A. 2013. Annotation Pro – a new software tool for annotation of linguistic and paralinguistic features. In Hirst, D. & Bigi, B. (eds.), Proc. of TRASP, 51-54, Aix-en-Provence, France.
- Malisz, Z. 2013. Speech rhythm variability in Polish and English: A study of interaction between rhythmic levels. Unpublished PhD thesis. Adam Mickiewicz University in Poznan.
- Yu J., Gibbon D. and Klessa K. 2014. Computational annotation-mining of syllable durations in speech varieties. In Campbell N., Gibbon D., Hirst D. (eds.), Proc. of the 7th Speech Prosody Conference, 443-447, Dublin, Ireland.

