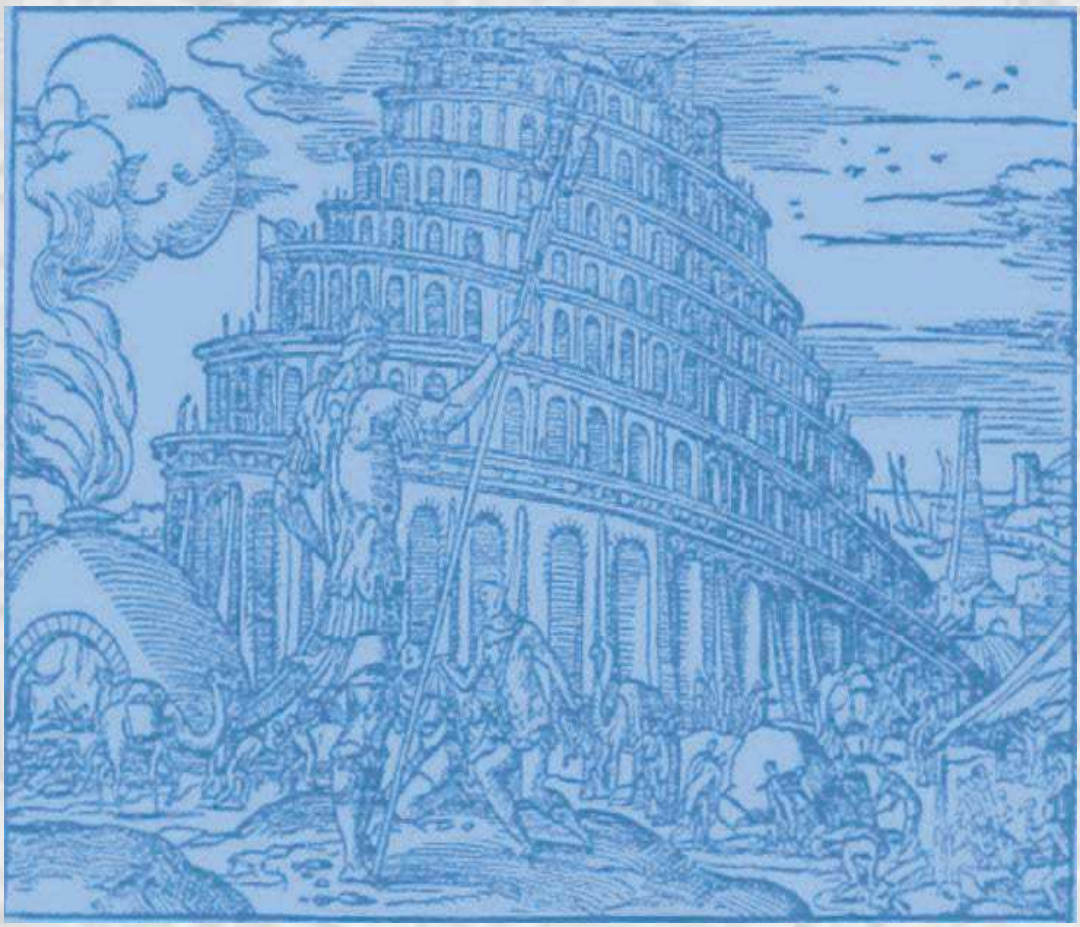
ExLing 2017

Proceedings of the 8th Tutorial & Research Workshop on

Experimental Linguistics

19-22 June 2017
Heraklion, Crete, Greece

Edited by Antonis Botinis



National and Kapodistrian
University of Athens

ExLing 2017

Proceedings of 8th Tutorial and Research Workshop on
Experimental Linguistics

19-22 June 2017, Heraklion, Crete, Greece

Edited by Antonis Botinis

National and Kapodistrian
University of Athens

ExLing 2017
Proceedings of 8th Tutorial and Research Workshop on Experimental Linguistics

Athens, Greece
Ebook ISSN: 2529-1092
Ebook ISBN: 978-960-466-162-6

Foreword

This volume includes the proceedings of ExLing 2017, the 8th Tutorial and Research Workshop on Experimental Linguistics, in Heraklion, Crete, 19-22 June 2017. The first conference was organised in Athens, Greece, in 2006, under the auspices of ISCA and is regularly repeated thereafter in different places.

In accordance with the spirit of the ExLing conferences, we were once again gathered in Crete 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 for a new generation of linguists. As in the 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 conferences, 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 invited speakers Kerstin Fischer, Philippe Martin, Yi Xu, and colleagues from the International Advisory Committee as well as our students from the University of Athens for their assistance.

Antonis Botinis

Contents

Tutorial papers

Research papers

Using robots to study speech – interactional processes and speech functions

Kerstin Fischer
Department of Design and Communication, University of Southern Denmark, Denmark.

Abstract

In this paper, I present some scenarios in which human-robot interaction can be carried out to study linguistic features and their interactional functions, as well as complex interactional processes.

Key words: speech, language, human-robot interaction, methodology

Introduction

Robots are a great methodological resource for linguistic and interactional research since they can be completely controlled (unlike human confederates) and thus exhibit identical behaviors across participants and conditions. Furthermore, people do not hold particular expectations towards robots, unlike towards other people, which allows us to experiment with a broad range of behaviors.

In the most trivial case, we can use robots in a variant of the matched guise technique (Gardner & Lambert 1972) to study people's attitudes towards certain linguistic features and their interpersonal functions; for instance, we can use them to study the functions of rising and falling intonation contours, speaking styles, durational features, the use of hesitation markers, types of feedback signals and so on. While the traditional matched guise experiment requires that there are several filler items between the two target stimuli that make people forget what they have just heard, in human-robot interaction, we can match the two sound files with different robots, or even easier, with two videos of different robots in a questionnaire. Besides this very simple and obvious use of robots in linguistic research, robots can also be a resource to study complex interactional mechanisms at the intersections between cognitive and social processes, which I will show below.

The next section presents some human-robot interaction scenarios in which linguistic research can be carried out. Then, I focus on one major interactional feature: recipient design, i.e. the processes involved in how we adjust our speech to our particular addressees, for instance, to a child, a foreigner or a robot. I illustrate how we can investigate complex socio-cognitive phenomena by means of human-robot interaction experiments.

Scenarios for linguistic human-robot interaction research

In the following, I describe some human-robot interaction scenarios that have already been successfully employed for linguistic research. While obviously many different scenarios are possible, the following are the most common ones.

Video-based studies

Experiments with robot videos are easy to set up and thus especially attractive for students; for instance, Aarstrup (2014) investigated how people respond to a greeting if it is delivered with different intonation contours. Previous linguistic work had suggested that, for instance, ``a fall indicates the speaker's dominance in knowing and telling something, in telling someone what to do, and in expressing their own feelings; in contrast, a rise indicates a speaker's deference to the addressee's knowledge, their right to decide, and their feelings" (Tench 1996: 105). Wells (2006) furthermore holds the greeting *'hi'* not to be used with rising intonation.

Aarstrup used videos of three different robots, in which the robots make a waving gesture or produce a small bow, and synthesized *hi* and *hello* with a free text-to-speech system. She then manipulated the sound files using praat (Boersma 2001) to create stimuli with different intonation contours. Altogether, the experiment uses two lexical items (*hello* vs. *hi*), three robots (Nao, Care-O-bot and Robosapien) and four different intonation contours (fall, rise, fall-rise and flat). The videos were then matched with the different sound files and distributed over four questionnaires so that each participant only got to see each robot once (n=120). Her results reveal that non-native speakers do not respond to the differences in intonation contours or lexical items at all. In contrast, native speakers of English rate the robots significantly different concerning friendliness, assertiveness and engagement depending on the intonation contours. However, these effects differ for the different lexical items, and the apparently non-conventional *hi* with rising intonation contour was in fact rated as most engaging.

Thus, matching videos with different speech stimuli provides a simple tool to investigate interpersonal effects of particular linguistic features.

One-on-one human-robot interactions

One-on-one interaction is the most common form of human-robot interaction experiments. The robot’s speech in this scenario can be pre-recorded, pre-synthesized or synthesized online using a text-to-speech system. The range of robots used in this scenario is principally without limits, but if the robot exhibits no human-like characteristics at all, speech may not be the most

plausible mode of communication. All speech phenomena that can be tested in video-based scenarios can also be studied in this scenario, plus most interactional features, such as the timing of utterances and the appropriateness of utterances in an interactional discourse context; in addition, properties of speech can be studied in relation to other modalities, such as gaze or robot movement, and their timing with respect to each other (e.g. Jensen et al., in press).

One human – two similar robots

In order to test behavioral effects of speech characteristics, Andrist et al. (2013, 2015) developed a scenario in which the participant is interacting with two similar robots. The two robots take turns presenting information to the participant; for instance, they both make suggestions for places to visit, from which the participant has to choose one. Crucially, the robots' speech differs with respect to the speech characteristic(s) under consideration. Andrist et al. (2013) investigate the effects of certain rhetorical strategies indicating expertise, Andrist et al. (2015) use the same scenario to study the interactions between clues to expertise and linguistic variety, in their case Standard Arabic versus the local Lebanese dialect. Their behavioral measure is how often participants follow the recommendations of the robot that used the strategy under consideration.

Using robots to investigate interaction processes

Recipient design is a central feature of all speech, but it becomes most obvious when people are talking to interaction partners who are 'peculiar' in some way or other, for instance, children, foreigners or dogs. While it is thus a common place that we adjust to our communication partners (cf. Brennan & Schober 2003), it is less clear how we do it.

In Fischer (2016), I carried out several human-robot interaction experiments to identify the contributions of a) a partner model, i.e. some kind of cognitive representation of the partner, of b) automatic adjustments to the partner's speech based on alignment (Pickering & Garrod 2004), and of c) the effects of the partner's feedback signals for the choices speakers make. For example, one set of experiments investigated how the way people talked to a robot changed depending on whether the robot produced speech itself. Because the linguistic output of the robot could be controlled completely, I could nail down in detail how much vocabulary and linguistic structures from the robots' utterances people aligned with and what else their linguistic behavior was informed by, disentangling the effects of partner models and automatic alignment. The results show that people's linguistic choices depend more on their partner model than on the re-use of the robot's linguistic behavior; in fact, the amount of 'automatic' alignment (priming)

was found to crucially depend on people's partner models. Similarly, the effects of the communication partner's (i.e. the robot's) contributions could be determined; while the robots' appearance was not found to have much influence, people use all kinds of behavioral clues to revise and update their partner models of the robots, which in turn determines their linguistic behavior.

To sum up, human-robot interactions are highly useful for all kinds of linguistic research, including the analysis of highly complex interactional phenomena.

Acknowledgements

The research reported on here was conducted within the project *Improving Second Language Pedagogy at the Prosody-Pragmatic Interface by Using Human-Robot Interaction*. The project is financed by the Danish Council for Independent Research.

References

Aarstrup, M.I.C. 2014. Greetings and Intonation Contours in Human-Robot Interaction. BA Thesis, University of Southern Denmark.

Andrist, S., Spannan, E., and Mutlu, B. (2013). Rhetorical Robots: Making Robots More Effective Speakers Using Linguistic Cues of Expertise. In *Proceedings of HRI '13*, pp. 341-348.

Andrist, S., Ziadee, M., Boukaram, H., Mutlu, B., and Sakr, M. (2015). Effects of Culture on the Credibility of Robot Speech: A Comparison between English and Arabic. In *Proceedings of HRI '15*, pp. 157-164.

Boersma, P. 2001. Praat: A system for doing phonetics by computer. *Glot International* 4, 341-345.

Fischer, K. 2016. Designing Speech for a Recipient. Partner Modeling, Alignment and feedback in so-Called Simplified Registers. Amsterdam: John Benjamins.

Gardner, R.G., Lambert, W. 1972. *Attitudes and motivation in a second-language learning*. Rowley Massachusetts: Newbury House Publishers.

Jensen, L.C., Fischer, K., Bodenhagen, L. Suvei, S.D. in press. The Synchronization of Multimodal Robot Behaviors during Human-Robot Collaboration. *Proceedings of Ro-Man'17*, Lisbon, Portugal.

Pickering, M.J., Garrod, S. 2004. Towards a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27, 169–225.

Schober, M.F., Brennan S.F. 2003. Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher, S. R. Goldman (Eds.), *Handbook of Discourse Processes*, pp. 123–164. Hillsdale: Lawrence Erlbaum.

Tench, P. 1996. *The Intonation Systems of English*. London: Cassell.

Wells, J. C. 2006. *English Intonation: an introduction*. Cambridge: Cambridge University Press.

Accent phrases and brain waves

Philippe Martin
LLF, UMR 7110, UFR Linguistique, Université Paris Diderot, France

Abstract

Recent experiments show that theta and delta brain oscillations whose periods vary from 100 ms to 250 ms and from 250 ms to 1300 ms are respectively synchronized in phase with syllables and stressed syllables. This means that, in natural speech, syllables cannot be separated by less than 100 ms and more than 250 ms to be perceived as unstressed. Over this limit, they are perceived as stressed but their separation cannot exceed 1300 ms. Considering the brain predictive function as observed in silent reading or self-talk, where restored syllables and stressed syllables activate theta and delta oscillations, a general language acquisition and speech recognition model is suggested.

Key words: Intonation, brain oscillations, delta, theta.

Introduction

Neurolinguistics is becoming one of the fastest changing domain in linguistics. After the syntactic and semantic domains, it now focuses on various aspects of speech generation, and in particular on specific speech events timing. Relationships between the perception of syllables, words or group of words boundaries and brain waves are now being investigated thoroughly in many research projects. In this contribution, I will evoke the intriguing similarities existing between stressed syllables timing and delta brain waves, leading to a model of speech processing where syllables, stressed syllables and accent phrases temporal properties are shown to correspond to those of theta, delta and gamma brain oscillations.

Brain waves

In the years 1930-40, researchers observed that the human brain consisted of a very large number of neurons (in the order of 100 billion) interconnected in groups in specific regions of the brain mass. These interconnections allow a transfer of chemically stored information in each neuron or groups of neurons by variations of a small electric potential (in the µV range). The electrical variations can be observed through captors positioned on subject's skull (electroencephalography, or EEG). These electrical variations are called evoked potential if they result from an external sensory stimulation, auditory, visual or other.

Electrical activity produced by transfers of group of neurons to other groups of neurons is not done haphazardly. First, they operate in specific

frequency ranges linked to specific cognitive activities, and secondly, they may synchronize in phase other waves in a different frequency range.

Greek letters designate specific frequency ranges: alpha, beta, delta, gamma. The frequency ranges of interest for speech are 1 Hz to 4 Hz for delta, 4 Hz to 10 Hz for theta, and 30 Hz to 80 Hz for gamma. The 12 Hz to 30 Hz range occupied by beta waves is also to consider. Since speech is essentially a temporal event, it may be more meaningful to characterize brain waves by their range of period variation, with delta periods varying from 250 ms to 1300 ms, theta 125 ms to 250 ms and gamma 12 ms to 33 ms.

The frequency of brain waves oscillations depends on the return distance from one neuronal region to another, so that the exact range of periods can vary from one subject to another. This transfer is done through a chain of synapses, whose individual length is about only 20 to 40 nm on the average.

Transfers between neuronal zones can be observed with a relatively large number of captors (from 32 to 256) placed around the subject skull according to location standards. Captured signals are stored in real time and analyzed into the frequency domain with either a (Fast) Fourier or Wavelet transform. The resulting representation is very similar to spectra obtained in frequency speech analysis, but in a lower frequency range.

Is then possible to observe that specific neuronal exchanges triggered by selected stimuli ("Event Related Potential", ERP) are phase coupled, with delta being the master. It has been shown for instance (Ghitza and Greenberg, 2009) that syllable peaks do synchronize theta brain waves in phase, whereas stressed syllable synchronize delta waves (Martin, 2015). Likewise, some specific frequency part of syllables, like fricative noise, synchronize gamma waves. In the absence of external stimuli, all oscillations oscillate freely in their own frequency range, delta with an at rest period of about 500 ms, and theta with an at rest period of about 125 ms.

Designing speech with available brain waves

Transfer of neuronal activity is essential for memorization, and some regions of the brain become specialized in information storage. Without memorization process, the image of an external sound stimulus is normally lost after 10 or 20 seconds. The acquisition of language requires thus some mechanism to encode and store the acoustic information provided by the speaker, for instance a mother to her child. The repetition of similar stimuli reinforces the neuronal memorization process, which has to be carried by some of the available brain waves.

As stressed syllables are the syllabic kernel of any word, or rather of any accent phrase (i.e. a group of syllables with one stressed syllables), a good candidate among brain waves would be delta, which would then phase lock theta oscillations while triggered by syllables (stressed or not). Gamma

oscillations have a supplementary phase lock role for the alignment of syllables.

Some experimental facts

It is relatively easy to consider theta as responsible for the syllabic information transfer. Slowing down natural speech without modifying its spectral content down to a 100 ms average syllabic duration will maintain its intelligibility. Going beyond this limit would make speech unintelligible, unless some silent gap is added to each syllable to reach the 100 ms average duration (Ghitza and Greenberg, 2009). Slowing down speech over a 250 ms average syllabic duration would on the other hand make every syllable perceived as stressed, although nothing has been changed in its spectral content.

These temporal limits suggest that to be perceived, syllables must have a duration over 100 ms and below 250 ms. Over 250 ms they are perceived as stressed. This strongly suggests that theta oscillations are carriers of non-stressed syllables. Over this 250 ms threshold, they will be handled by delta oscillations.

Likewise, it is relatively simple to show that stressed syllables must be at least separated by 250 ms by selecting two consecutive stressed syllables in natural speech and by reducing their gap, normally filled by a pause, without modifying their relative spectra. Once their separation goes below 250 ms the first syllable ceased to be perceived as stressed. On the other hand, two consecutive stressed syllables appear to be impossible to separate in natural speech by more than some 1300 ms (Martin, 2014). These two observations suggest that delta is the candidate carrier for stressed syllables.

Building the lexicon

A possible model for early language acquisition would then involve the elaboration in early childhood of a lexicon whose entries are activated by delta waves triggered by stressed syllables (including isolated syllables). Isolated vowels would progressively by replaced by full syllables, then by groups of syllables with one stressed syllable (i.e. an accent phrase), then later by sequences of accent phrases forming sentences, and eventually larger speech units involving more than one sentences (when leaning poems by hart for instance).

The various units in this lexicon are used to identify incoming speech by comparison with an existing entry, at one or more appropriate levels (syllabic, accent phrase, syntagms, etc.). Delta is the master of this continuous pattern matching operation, which requires at least 250 ms to be

executed. If two stressed syllables occur at less than 250 ms apart, the first delta cycle is aborted and only the second one is executed.

Predictive function of the brain

The pattern matching model gives also some account for the predictive function of the brain (Arnal et Giraud, 2016). Listeners sometimes perceive stressed syllables although nothing in the speech signal carries the corresponding acoustic information. This happens when the pattern matching process found an entry in the listener lexicon that would reasonable match the speech input, restoring the stress information missing in the speech signal.

It gives also a reasonable explanation pertaining to the continuous adjustment of the delta periods according to the stressed syllables speech input. The natural oscillation normally cycles by retaking the last period value for the next, a process whose effects can be observed in eurhythmicity (Martin, 2015). However, when a stressed syllable occurs before the end of the expected period, the delta oscillation phase is reset and a new period starts. When it occurs after, the predicted period duration inferred from the lexicon retrieved entry will set the start of the new period.

This may explain why theta and delta oscillations are restored even in the absence of any speech input when reading or taking to oneself silently, their respective cycle duration being restored from the lexicon.

References

Arnal, L.H., Giraud, A-L., M. 2016. Neurophysiologie de la perception de la parole et multisensorialité, in Traité de neurolinguistique, De Boeck, Louvain-la-Neuve, 416 p.

Ghitza O" Greenberg S. 2009. On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence, Phonetica. 2009, 66(1-2), 113-26.

Martin, Ph. 2014. Spontaneous speech corpus data validates prosodic constraints, Proceedings of the 6th conference on speech prosody. Campbell, Gibbon, and Hirst (eds.), 525-529.

Martin, Ph. 2015. The Structure of Spoken Language. Cambridge University Press.

Syllable as a synchronization mechanism

Yi Xu
Dept of Speech, Hearing and Phonetic Sciences, University College London, UK

Abstract

Despite being highly intuitive and widely recognized, syllable continues to be a controversial notion. It is argued here that a resolution may lie in recognizing that speech is a highly skilled motor activity with a core problem shared with other motor skills: how to reduce degrees of freedom (DOF) to the extent that makes its central nervous control possible. The most effective way of reducing DOF is to *synchronize* multiple articulatory movements, and the syllable serves exactly this function for speech. This synchronization hypothesis also offers resolutions to coarticulation and many other unsettled problems, and has implications for motor control in general.

Key words: synchronization, target approximation, edge synchronization, tactile anchoring, degrees of freedom

Introduction

The nature of the syllable remains a mystery to this day: "*although nearly everyone can identify syllables, almost nobody can define them*" (Ladefoged 1982: 220). Existing proposals (e.g., Browman & Goldstein 1992, MacNeilage 1998) are not able to answer some of the most basic questions about the syllable: a) *Why are there syllables*? b) *Do syllables have clear phonetic boundaries*? c) *Do segments have definitive syllable affiliations*? Without clear answers to these questions, many other issues about speech also remain unresolvable, including, in particular, coarticulation.

Syllable as a synchronization mechanism

This article is a brief introduction of a *synchronization hypothesis* that can address all three questions mentioned above. The overarching proposal is that syllable is a temporal coordination mechanism whose function is to synchronize multiple articulatory movements so as to make speaking possible. The coordination involves three basic mechanisms: *target approximation*, *edge synchronization* and *tactile anchoring*.According to this hypothesis, speech encodes information by generating variations in phonetic (segmental, tonal and phonational) properties in quick succession, which requires concurrent articulatory movements toward multiple underlying targets (*target approximation*). The central nervous control of the concurrent movements is made possible by synchronizing the onsets and offsets of the movements (*edge synchronization*) to critically reduce degrees of freedom

(DOF). And tactile sensation during the closed phase of each syllable provides alignment references for the synchronization of movements (*tactile anchoring*).

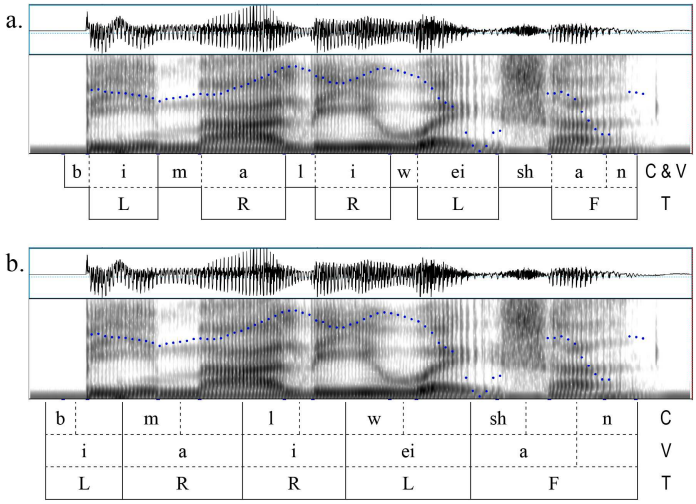


Figure 1. Spectrogram of the Mandarin phrase "比麻黎偽善" /bǐ má lǐ wěi shàn/ [more hypocritical than Ma Li], with broad phonetic transcriptions. In both panels, C, V and T stands for consonant, vowel and tone. The segmentation in a. is conventional, while that in b. is based on the synchronization hypothesis.

A direct consequence of synchronization hypothesis is a major change in the acoustic segmentation of the syllable, as illustrated in Figure 1. The conventional segmentation is shown in Figure 1a, where syllable onsets are aligned to points of abrupt spectral shift corresponding to the moments of complete oral closure (with the exception of /w/, for which there is not even a widely-agreed segmentation). The segmentation based on the synchronization hypothesis is shown in Figure 1b, where all the segmental boundaries are shifted leftward from those in Figure 1a. These *conceptual* shifts are based on the *target approximation* principle that the onset of a segment is when the spectral pattern *starts to move toward* its prototypical configuration. For /l/, for example, the onset is in the middle of the "preceding" vowel, where F1 starts to drop toward the oral closure. The leftward shift of vowel onsets is even more extensive. For example, the first /a/ now begins from the middle of the first conventional /i/ interval where F2 starts to drop toward its prototypical level; and the second /i/ begins from the middle of the conventional /a/ interval where F2 starts to rise toward its prototypical level. Thus each vowel onset is shifted leftward, across the

conventional consonant interval, and well into the conventional interval of the preceding vowel.

The new *offset* of a segment is where its prototypical configuration is best approach, but not necessarily attained. For /i/, for example, it is at the peak of F2 and F3, for /a/ it is at the peak of F1 and valley of F2, and for /w/ it is at the valley of F2, where intensity is also at a minimum. For the obstruent consonants, the offset is no longer at the end of its prototypical spectral pattern (e.g., closure gap in /b/, nasal or lateral formants in /m/ and /l/, and the frication in /ʂ/), but in the middle of those intervals.Furthermore, as shown in the bottom tier of Figure 1b, the new tone interval fully coincides with that of the entire syllable, which consists of not only the initial consonant and the nuclear vowel, but also the coda consonant, as in the case of /shàn/.

Coarticulation and DOF

A major impact of the new syllable segmentation is on the understanding of coarticulation. Because the initial movement toward a vocalic target is now viewed as the vowel proper rather than its anticipation, it is no longer considered as evidence of either long-distance anticipatory V-to-V coarticulation (Öhman 1966) or local V-to-C anticipatory coarticulation. Instead, because initial consonant and the first vowel start at the same time, they are considered as fully *coproduced* or coarticulated for the duration of the consonant. Also no longer needed is the notion of carryover coarticulation. Due to inertia, an articulatory gesture toward a target has to be one of moving away from its initial state, which is the end result of approaching the preceding target. Overcoming inertia therefore necessarily carries the influence of the preceding target that is no longer being executed. So, what is carried over is only its remnant effect rather than its continued articulation.

This view of coarticulation differs from both the assimilation account and gestural overlap account (Saltzman & Munhall 1989), as illustrated in Figure 2. In panel *a* we can see that smooth surface trajectories can be generated by strictly sequential target approximation movements. In panel *b* the first movement is much shortened from that in panel *a*, resulting in an undershoot of the first target. Because the undershoot is due to a premature termination of the first movement, which is effectively *truncated* by the second movement, there is neither *assimilation* nor *anticipation*. In panel *c*, instead of truncation, the first two movements are partially overlapped via *gestural blending*. The resulting trajectory, however, is not very different from the one in panel *b*. More importantly, such a blending requires *more degrees of freedom* than sequential target approximation, as the amount of overlap and its exact location both have to be specified.

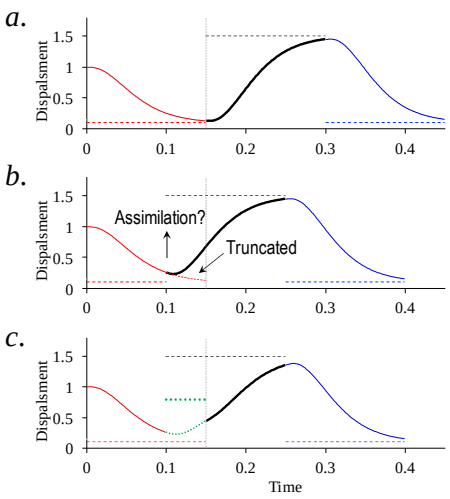


Figure 2. Sequential vs. blending target approximation. In *a*, the three movements are strictly sequential, and the vertical line divides the first two movements. In *b*, the vertical reference remains at time 0.15, but the first movement is shortened by 0.05 unit. All the movements remain sequential, so that the first movement is *truncated* by the second. In *c*, the first and second movements overlap with each other by 0.05 units. The overlap is implemented by applying a blended target (horizontal green dotted line), which is the average of the first two targets. All trajectories generated by the qTA model (Prom-on et al. 2009).

Tactile anchoring

This mechanism is the final piece that completes the syllable puzzle, because it provides solutions to three critical issues. First, it explains how synchronization is achieved: by tactile feedback during the closure phase of consonants. Second, it points out that it is the edges, rather than the center of the syllable (where sonority is the highest, cf. review by Ohala 1992), that is the essence of the syllable. Finally, it predicts that syllable onset, where a maximum number of gestures can be coproduced, is a better synchronization site than syllable offset, which in turn explains why CV is the more prevalent syllable structure than VC and CVC.

References

Browman, C.P., Goldstein, L. 1992. Articulatory phonology. Phonetica 49, 155-180.

MacNeilage, P.F. 1998. The frame/content theory of evolution of speech production. Behavioral and Brain Sciences 21, 499–546.

Mechsner, F., Kerzel, D., Knoblich, G., Prinz, W., 2001. Perceptual basis of bimanual coordination. *Nature* 414: 69-73.

Ladefoged, P. l982. A Course in Phonetics. University of California, Los Angeles.

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

Prom-on, S., Thipakorn, B., Xu, Y. 2009. Modeling tone and intonation in Mandarin and English as a process of target approximation. JASA 125.

Saltzman, E.L., Munhall, K.G. 1989. A dynamical approach to gestural patterning in speech production. Ecological Psychology 1, 333-382.

Loanwords stress and intonation in Algerian Arabic

Ismaël Benali
CLILLAC-ARP, Université Paris Diderot, France

Abstract

This paper deals with stress and intonation patterns of loanwords in Algerian Arabic and in particular in Algiers and Oran dialects. It shows that speakers always transferred the stress pattern of Algerian Arabic in borrowed words: the final syllable is stressed if it is the only heavy syllable in the word or if it is superheavy, otherwise stress is on the penultimate syllable. Intonation patterns are also transferred to loanwords in focus and polar questions. Focus is characterized in the Algiers dialect by a falling melodic contour occurring on the last syllable. In the Oran dialect, intonation is realized with a flat or slightly rising contour. In polar question Algiers speakers produce a rising-falling contour whereas the Oran speakers produce a rising melodic contour on the last syllable.

Key words: loanword, Algerian Arabic, stress, focus, polar question

Introduction

The Algerian Arabic originates from the North African Arabic and is spoken by more than 70% of the population. It is the main vehicular language of the country. The Algiers variety is spoken in the capital in the Center of Algeria. The Oran variety is spoken in the city of Oran in the West of Algeria. In previous studies (Benali, 2008, Benali, 2016), it appeared that Algerian dialects can be identified by prosodic cues. I found also that intonation patterns which characterize Algiers and Oran varieties are marked more clearly in focus and polar question.

These two varieties have great number of borrowed words from French whereas the Oran dialect contains also Spanish loanwords. These loanwords are phonetically and morphologically modified (Gallicism) and adapted to the speakers mother tongue (Arabic or Berber) as in [si.'ni.ma] from the French word "cinéma" and in [bo.'ga.do] from the Spanish word "abogado".

In this research, I examine the place of the stress of loanwords and their prosodic realization in focus and polar question.

Few studies on stress were carried out on Algerian Arabic. Aït Oumaziane Ramadane studied the stress in the Arabic dialect of Constantine (Aït Oumeziane, 1981), Farouk Bouhadiba in the Oran dialect (Bouhadiba, 1988) and Aziza Boucherit in Algiers Arabic (Boucherit, 2006). From our study (Benali, 2015) stress falls on the final syllable if it is the only heavy syllable

in the word (V: or VC), or if it is superheavy V:C(C) or VCC. Otherwise the penultimate syllable is stressed.

Few studies have dealt with the integration of loanwords into dialectal Arabic at the prosodic level. Borrowed words from French tends to lose their " tonic" accent to submit to the accentual rules of the Tunisian dialectal Arabic (Mzoughi, 2015). Naïma Fadil-Barillot presented three types of the intonative contours in the code-switching of Morroccan Arabic and French : the intonation of matrix language, the composed intonation and the recomposed intonation (Barillot, 2002).

Methodology

Recordings for this study were made in Algiers and Oran. 20 Algiers speakers and 20 Oran speakers were recorded in a quiet room using a lapel microphone. Speakers were between 22 and 30 years old. The majority of them were students and have lived all their childhood and adolescence in their respective cities.

The corpus is composed of spontaneous and read speech. The spontaneous speech utterances were extracted from a conversation between speakers and the experimenter (who was speaking a southern variety). Borrowed words and code switching were observed in this corpus.

Acoustic analysis were carried on the speech analysis/resynthesis program 'WinPitch' (Martin, 2000).

Analysis

Stress in spontaneous speech

The use of the loanword "ball" in our corpus illustrates the difference between the two dialects: Algiers speakers use the French word "ballon" and Oran speakers use the Spanish one "bola". In both cases stress is on the penultimate: [ˈba.lɔ̃], [ˈbo.la].

Overall, Oran speakers stress the penultimate syllable as in [ˈsba.ɲa] "Spain" and the final syllable when heavy and closed: [sba.ˈɲo:l] "Spanish". In the word Cervantes the stress is not on the penultimate as in Spanish but in the last syllable.

The following example shows how the loanword is morphologically integrated: [ˈkunt ˈ**ntri.ni** u ħa.ˈbəst] "I trained and then I stopped" [ˈntri.ni] from the French verb "entraîner", [n] and [i] are Arabic morphological marks. Stressed syllable is longer and has a rising F0.

Algiers speakers stress also the penultimate syllable as in [teks.ˈplo.zi] "you explode (in anger)" from the French verb "exploser". This syllable is lengthened whereas the finale syllable carries a falling F0 as in [ˈdʒa.zət

in.ˈ**ti:k**] "It went very well" from the French word "antique" in the meaning of well, good.

So Algiers and Oran speakers in loanwords stress the final syllable if it is the only heavy syllable in the word or if it is superheavy. In the other case the penultimate is stressed. Acoustic analysis shows that Algiers speakers used more F0 range to stress the syllable than the Oran speakers ($p < 0{,}0001$) who use more vowel lengthening ($p < 0, 0013$).

Intonation in read speech

In the statement used for this study [ˈʃra kat.sã.ˈkat ˈba.ʃi] "He bought a 404 pickup", there are two loanwords [kat.sã.ˈkat] from the French model car "404 Peugeot" and [ˈba.ʃi] from the French word "bachée" meaning a pickup. Only the second word is stressed on the penultimate syllable, forming an accent group. The speakers were asked to produce this statement with an emphatic focus on the last word and to produce it as a polar question.

Emphatic narrow focus is produced in the Algiers dialect by a rising contour on the stressed syllable followed by a falling contour. In the Oran dialect, this focus is realized with a flat or slightly rising contour on the stressed syllable (figure 1). In both dialects the stressed syllable is lengthened.

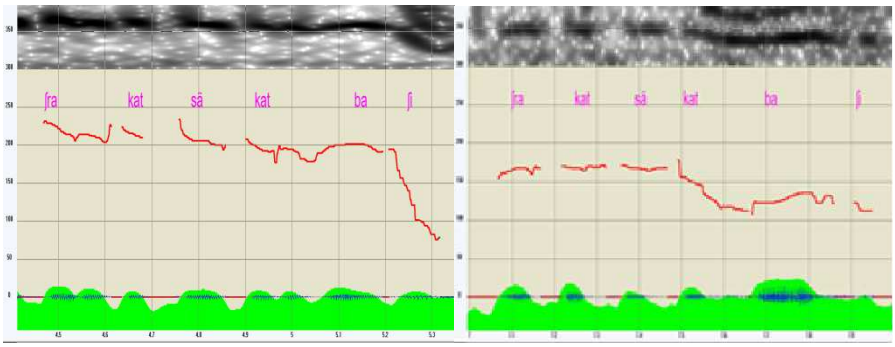


Figure 1. Emphatic narrow focus produced by Algiers (left) and Oran speakers (right).

In polar question Algiers speakers produced an amplified rising-falling contour while Oran speakers produced on the last syllable a rising contour (figure 2).

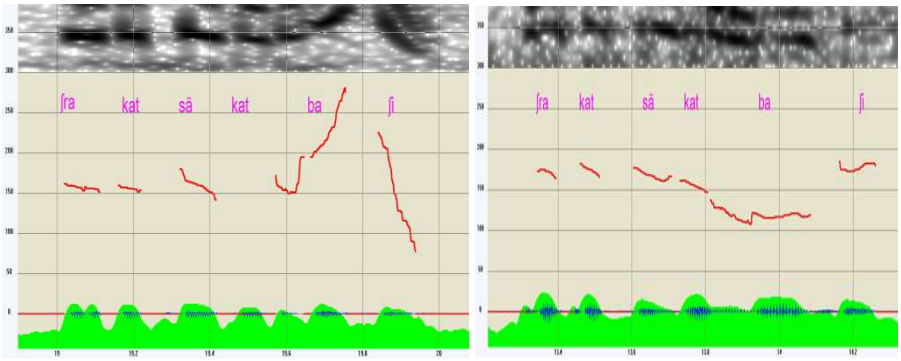


Figure 2 Polar question produced by Algiers (left) and Oran speakers (right).

These results are similar to those obtained in Algiers and Oran dialectal Arabic.

Conclusion

The results of this study show that Algerian speakers always transferred the stress pattern of Algerian Arabic to loanwords words. Intonation patterns are also transferred in focus and polar questions.

References

Aït Oumeziane, R. 1981. *Le Parler Arabe de Constantine*. Doctorat, Université Sorbonne Nouvelle Paris 3.

Barillot, N.F. 2002. *Etude sociolinguistique et intonative de l'alternance codique arabe marocain français.*

Benali, I. 2008. La question totale dans les parlers algérois et oranais. *Language design: journal of theoretical and experimental linguistics***,** 0039-46.

Benali, I. 2015. *La focalisation dans les parlers algérois et oranais: étude prosodique*. Sorbonne Paris Cité.

Benali, I. 2016. The identification of two Algerian Arabic dialects by prosodic focus. *ExLing 2016***,** 37.

Boucherit, A. 2006. Algiers Arabic. *Encyclopedia of Arabic Language and Linguistics***,** 58-66.

Bouhadiba, F.A.N. 1988. *Aspects of Algerian Arabic verb phonology and morphology*. University of Reading.

Martin, Ph. WinPitch 2000: a tool for experimental phonology and intonation research. Proceedings of the Prosody 2000 Workshop, 2000.

Mzoughi, I. 2015. *Intégration des emprunts lexicaux au français en arabe dialectal tunisien*. Université de Cergy Pontoise.

Modelling prosodic structure using Artificial Neural Networks

Jean-Philippe Bernardy, Charalambos Themistocleous
CLASP, University of Gothenburg, Sweden

Abstract

The ability to accurately perceive whether a speaker is asking a question or is making a statement is crucial for any successful interaction. However, learning and classifying tonal patterns has been a challenging task for automatic speech recognition and for models of tonal representation, as tonal contours are characterized by significant variation. This paper provides a classification model of Cypriot Greek questions and statements. We evaluate two state-of-the-art network architectures: a Long Short-Term Memory (LSTM) network and a convolutional network (ConvNet). The ConvNet outperforms the LSTM in the classification task and exhibited an excellent performance with 95% classification accuracy.

Key words: Cypriot Greek, statements, questions, convolutional networks, LSTMs.

Introduction

The aim of an intonational model is to identify and represent the tonal structure of an utterance for the purposes of classification or generation of tonal contours. Each sentence is characterized by distinct *F*0 patterns that are perceived as different melodies, such as questions, statements, commands, and requests. The purpose of this study is to provide a classification model of Cypriot Greek statements and questions, using information from their fundamental frequency (F0) contours (for a description of Cypriot Greek intonation see Themistocleous, 2011, 2016). To this purpose two different neural network architectures were tested: a Long short-term memory (LSTM) neural network and a convolutional neural network (ConvNet).

Properties of tonal contours

In statements with broad focus, each phonological word associates with rising and falling tunes; the whole melody starts usually at a higher frequency and declines gradually to a lower frequency (see Figure 1 upper right panel). By contrast, Wh-questions are characterized either by sustained tune at a high frequency (see Figure 1 upper left panel) or by high-low tune at the Wh-word followed by a fall and an optional rise at the end of the utterance (Figure 1 lower panel). Notably, the automatic classification of intonation is challenging as tonal contours can differ depending on linguistic and non-linguistic factors including the segmental structure of utterances, the distribution of lexical stresses, the speech rate, and the physiology of

speakers (gender, age, etc.). These factors affect the overall structure of tonal patterns, such as the frequency range and the length of the utterance.

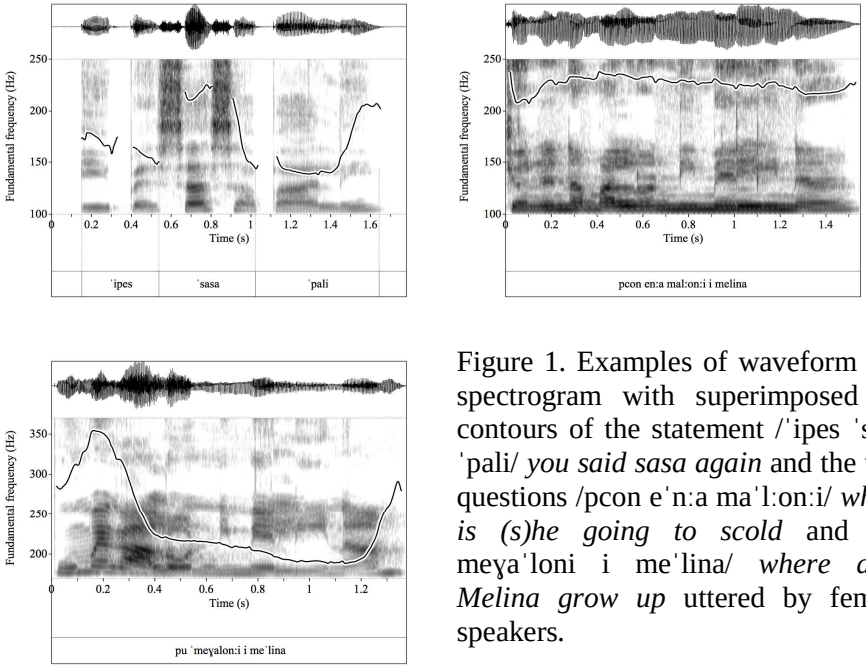


Figure 1. Examples of waveform and spectrogram with superimposed F0 contours of the statement /ˈipes ˈsasa ˈpali/ *you said sasa again* and the wh-questions /pcon eˈnːa maˈlːonːi/ *whom is (s)he going to scold* and /pu meɣaˈloni i meˈlina/ *where does Melina grow up* uttered by female speakers.

Artificial network architectures

LSTMs are recurrent neural architectures (RNN) (Hochreiter and Schmidhuber, 1997), which proved to be extremely powerful and dynamical systems in recognizing, processing, and predicting tasks such as time series, which are characterized by time lags of unknown size and bound between important events (Schmidhuber, Wierstra and Gomez, 2005). LSTMs displayed a strong record for addressing problems in speech recognition (Graves, et al. 2013), rhythm learning (Gers, Schraudolph, and Schmidhuber, 2002) etc. However, a disadvantage of using LSTMs (and RNNs in general) is that they are slower with respect to feed-forward networks; such as convolutional neural networks (ConvNets). ConvNets are feed-forward artificial neural networks, which learn patterns in restricted regions a.k.a., receptive fields that partially overlap. These architectures were inspired by biological processes of visual perception. These architectures were successfully employed to address tasks in image recognition (Ciresan, Meier, and Schmidhuber, 2012) and sentence processing.

Experiments

For the purposes of this study, 1966 statement productions were produced by 25 female speakers and 2860 Wh-questions were produced by 20 female speakers of Cypriot Greek. Sound files were segmented and labelled and the corresponding F0 contours were extracted every 12.5 ms. The data was then padded with zeros for a total length of 1024 and provided as an input to our models. Statements consisted of the same utterance melody "ipes CVCV pali" *you said CVCV again*, where C stands for a consonant and V for a vowel whereas Wh questions varied both in length and in the initial Wh-word. For the purposes of this study a twofold experiment has been employed, namely a classification task with an LSTM neural network and a classification task with a ConvNet.

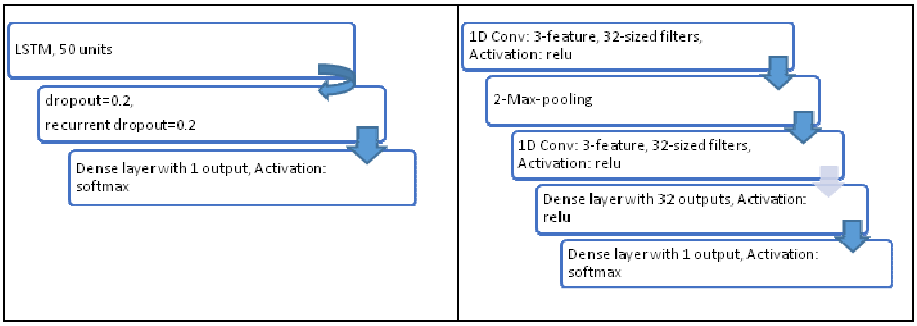


Figure 2. LSTM Neural Network Architecture (left panel) and ConvNet Architecture (right panel).

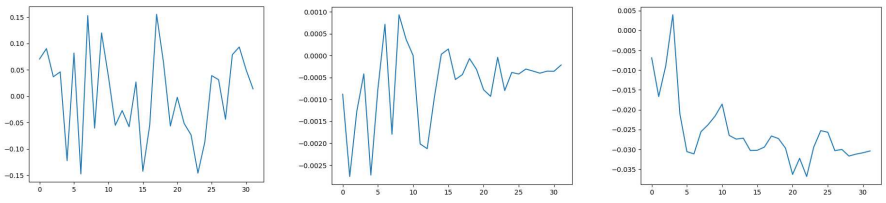


Figure 3. Examples of learned convolutional filters for the first layer.

For each run models were trained for 18 epochs with backpropagation. After each epoch, a model is obtained. Out those 18 models, the model yielding the smallest loss on the training data is kept, and used for evaluation against the test set. The LSTM resulted in 82% classification accuracy, while the ConvNets resulted in over 95% accuracy. For the ConvNet, we performed 10-fold cross validation (4343 training examples and 434 test

examples). The distribution of the classification accuracy has median 95%. (Min. 90%, 1st Qu. 94%, Median 95%, 3rd Qu. 96% and Max. 97%).

Discussion

We have described two artificial neural network architectures: an LSTM and a convolutional network. The ConvNet outperformed the LSTM in accuracy and speed. Specifically, the ConvNet achieves high performance on question and statement classification without requiring external features, such as manually tagged pitch accents. The network can identify by itself the properties of the tonal contour that are most significant for the classification. Despite the high accuracy learned features are not easily interpreted from a linguistic point of view (see Figure 3), yet a promising aspect of this architecture is that it requires few parameters (only 6 filters). However, this study was performed on a corpus that exhibits regularities and it remains to be seen whether it generalizes in a more varied corpus.

References

Ciresan, D., Meier, U., Schmidhuber, J. 2012. Multi-column deep neural networks for image classification. *IEEE Conference on Computer Vision and Pattern Recognition.* New York, NY: Institute of Electrical and Electronics Engineers (IEEE): 3642–3649.

Felix A. Gers, Jürgen Schmidhuber, Cummins Fred. 2000. Learning to Forget: Continual Prediction with LSTM. *Neural Computation.* 12 (10): 2451–2471.

Graves, A., Mohamed, Abdel-rahman, Hinton, G. 2013. Speech Recognition with Deep Recurrent Neural Networks. *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*: 6645–6649.

Gers, F. Schraudolph, N., Schmidhuber, J. 2002. Learning precise timing with LSTM recurrent networks. *Journal of Machine Learning Research*. 3: 115–143.

Hochreiter, S., Schmidhuber, J. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.

Kalchbrenner, N., Grefenstette, E., Blunsom, Ph. 2014. A Convolutional Neural Network for Modelling Sentences. arXiv:1404.2188

Schmidhuber, J., Wierstra D., Gomez F.J. 2005. Evolino: Hybrid Neuroevolution / Optimal Linear Search for Sequence Learning. *Proc. of the 19th International Joint Conference on Artificial Intelligence (IJCAI).* Edinburgh, 853–858.

Themistocleous Charalambos 2011. Prosody and Information Structure in Greek (in Greek). PhD Thesis. University of Athens Greece.

Themistocleous, Charalambos 2016. Seeking an anchorage: Evidence from the tonal alignment of the Cypriot Greek prenuclear pitch accent. *Language and Speech*., 59(4). 433-461, doi: 10.1177/0023830915614602.

Syllable phonology and constituency temporal production in Greek

Antonis Botinis[1], Evgenia Magoula[2], Olga Nikolaenkova[3], Athina Tsiori[2]
[1]Lab of Phonetics & Computational Linguistics, University of Athens, Greece
[2]Department of Primary Education, University of Athens, Greece
[3]Department of General Linguistics, Saint Petersburg State University, Russia

Abstract

This is an experimental study of temporal organisation of the Greek syllable. In accordance with a production experiment, the results indicate the following: (1) open and closed syllable units have a significant difference. (2) open and closed syllables do not have significant differences of vowel nucleus durations. (3) onset syllable consonants are significantly longer than coda syllable consonants. (4) lexical stress application has a significant lengthening effect on syllable unit, onset consonant as well as nucleus vowel but not on coda consonant constituents. (5) focus application does not have any significant effect on any syllable constituent.

Key words: consonant, vowel, duration, syllable, temporal production, stress, Greek

Introduction

The present study is an experimental investigation of segment durations as a function of syllable structure, lexical stress and focus. Thus, the main questions concern (1) the effects of each of the above prosodic factors on segment durations and (2) the interactions among the above factors. Syllable structure involves reverse phonotactis, i.e. CV and VC, and thus open vs. closed syllable structures in variable syllable unit contexts.

In accordance with research in different languages, including Greek, a variety of hypotheses with reference to segment duration variability has been suggested. Among them, more consonants in syllable onset are correlated with shorter respective durations (Botinis, Erkenborn, Isacsson, Westin, 1999), open syllable structure is correlated with longer vowel nucleus than closed syllable structure (Maddieson 1985) and stressed syllables are correlated with longer consonant onset as well as vowel nucleus (Botinis 1989, Fourakis, Botinis, Katsaiti 1999).

However, despite significant research, segment temporality as a function of syllable constituency variability has hardly been investigated. E.g., although it is widely known that lexical stress has a lengthening effect at syllable level, the effects of lexical stress on different syllable constituents are hardly known. Thus, in this paper, we attempt to enlarge our knowledge on temporal correlations as a function of syllable constituency variability.

Experimental methodology

The speech material consists of four test words: two in nominative singular and two in accusative plural with lexical stress at the antepenultimate and penultimate, respectively (Table 1). The test words were produced at the beginning of the carrier phrase [___ˈfonakse ðinaˈta] ‘s/he shouted ___ loudly’. Five female students at their mid-twenties, with standard Athenian Greek pronunciation, produced the speech material at a normal tempo in focus and out of focus context in a sound-treated studio at Athens University Phonetics laboratory. The speech material was analysed with Praat programme and segment duration results were subjected to statistical processing with SPSS statistical package.

Table 1. Test words in nominative and accusative with lexical stress assignment in antepenultimate and penultimate syllable, respectively.

Nominative singular	Accusative plural	Gloss
ˈenɵetos	enˈɵetus	Inserted
ˈnefelos	neˈfelus	Nefelos (name)

Results

The results are shown in figures 1-3. In accordance with a three-way ANOVA (syllable type x lexical stress x focus), syllable type and lexical stress have significant effects on both onset consonant and vowel nucleus durations whereas focus has no significant effect on any syllable constituent.

Figure 1 shows mean durations of syllable type units as well as syllable constituents as a function of open vs. closed syllables. The open syllable unit is 187 ms (SD 52) and the closed syllable unit is 170 ms (SD 45) and this difference of 17 ms is significant ($F=7.1$, $p<0.008$). The onset consonant in open syllable is 87 ms (SD 26) whereas the coda consonant in closed syllable is 58 ms (SD 11), a significant difference of 29 ms ($p<0.0001$). The nucleus vowel in open syllable is 100 ms (SD 32) and in closed syllable 111 ms (SD 42), a non-significant difference of 11 ms.

Figure 2 shows mean durations of syllable type units as well as syllable constituents as a function of lexical stress application. The stressed syllable is 211 ms (SD 41) and the unstressed syllable 145 ms (SD 31), a significant difference of 66 ms ($F=94$, $p<0.0001$). The onset consonant in stressed syllable is 101 ms (SD 23) and in unstressed syllable 71 ms (SD 20), a significant difference of 30 ms ($F=27$, $p<0.0001$). The vowel nucleus in stressed syllable is 131 ms (SD 28) and in unstressed syllable 79 ms (SD 27), a significant difference of 52 ms ($F=100$, $p<0.0001$. The coda consonant in

stressed syllable is 57 ms (SD 14) and in unstressed syllable 59 ms (SD 7.5), which is a non-significant difference of 2 ms.

Figure 3 shows mean durations of syllable constituents as a function of focus application. The effects of focus application are in general negligible and do not reach a significant level on either syllable unit or any onset, nucleus or coda syllable constituent.

The results indicate significant interactions of syllable type x lexical stress with reference to syllable unit ($F=14.2$, $p<0.0003$) but not any other syllable constituent.

In accordance with the above results, lexical stress application has a bigger temporal effect on syllable unit and/or syllable constituents than syllable type whereas focus application hardly has any effect. The temporal effect of lexical stress application, on the other hand, has a hierarchical effect on different syllable constituents, i.e. nucleus vowel>onset consonant>coda consonant.

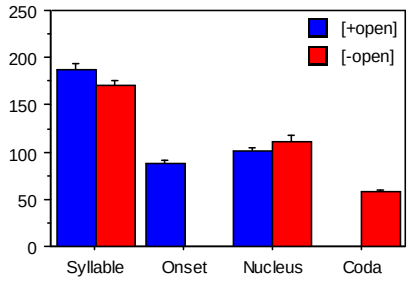


Figure 1. Syllable unit as well as onset, nucleus and coda constituent durations (in ms) as a function of open ON (+open) vs. closed NC (-open) syllable structure.

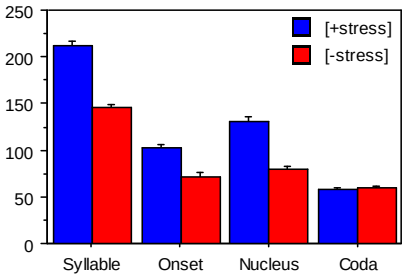


Figure 2. Syllable unit as well as onset, nucleus and coda constituent durations (in ms) as a function of stressed (+stress) vs. unstressed (-stress) syllables.

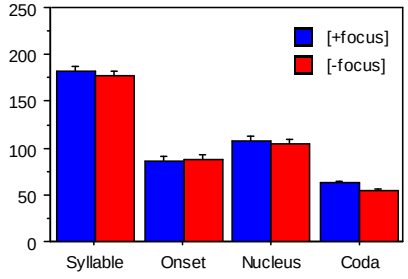


Figure 3. Syllable unit as well as onset, nucleus and coda constituent durations (in ms) as a function of focus (+focus) vs. out of focus (-focus) context.

Discussion and conclusions

This study is an investigation of temporal organization of syllable structure, i.e. VC vs. CV, as a function of lexical stress and focus applications. The main results indicate the following: (1) CV syllable is longer than VC syllable and this is due to the onset consonant of open syllable rather than its nucleus vowel. (2) onset consonant of CV syllable is longer than coda consonant of VC syllable. (3) lexical stress application has a lengthening effect on onset consonant and nucleus vowel but not on coda consonant. (4) focus application has no temporal effect on any syllable constituent.

The results of the present study are hardly in accordance with most studies in syllable structure and duration correlates. Most importantly, there is no evidence of the open syllabicity lengthening effect of nucleus vowel, as suggested by Maddieson (1985) and others (see e.g. Farnetani, Kori 1986, McCrary 2004). On the other hand, similar to the open syllabicity lengthening effect is evident in another study (Chaida et al. 2017, this volume). However, in the latter study, there was a difference speech material, i.e. CV vs. CVC, and hence different syllable structure. Furthermore, there was a compensatory lenthening effect, according to which longer nucleus vowel entailed shorter onset consonant. Its seems that the temporal organisation of syllable is the result of many factors among which the open syllable lengthenis effect is one among them.

Acknowledgements

Our thanks to Yi Xu for comments and much useful feedback as well as Ilectra Dimoula and Maria Papanagiotou for data recordings and data analysis. Thanks also to Special Account for Research Grants and National and Kapodistrian University of Athens for research grant and travel allowance.

References

Botinis, A. 1989. Stress and Prosodic Structure in Greek. Lund University Press.

Botinis, A., Erkenborn, S., Isacsson, C., Westin, P. 1999. Prosodic variability and segmental durations in Greek and Swedish. Proc. VIIth Swedish Phonetics Conference FONETIK 99. Gothenburg Papers in Theoretical Linguistics 81, 27-30.

Chaida, A., Dimoula, I., Magoula, E. Nikolaenkova, O. 2017. Open vs. closed syllable phonology and temporal production in Greek (this volume).

Farnetani, E. and Kori, S. 1986. Effects of syllable and word structure on segmental durations of spoken Italian. Speech Communication 5, 17-34.

Fourakis, M., Botinis, A., Katsaiti, M. 1999. Acoustic characteristics of Greek vowels. Phonetica 56, 28-43.

Maddieson, I. 1985. Phonetic cues to syllabification. In Fromkin, V. (ed.), Phonetic Linguistics, 203-221. London: Academic Press.

McCrary, K.M. 2004. Reassessing the role of the syllable in Italian phonology. Ph.D. thesis, UCLA.

The entrainment of EEG delta oscillations in speech listening

Victor J. Boucher[1], Annie C. Gilbert[2], Boutheina Jemel[3]
[1] Département de linguistique, Université de Montréal, Canada
[2] School of Communication Sciences and Disorders, McGill University, Canada
[3] Département d'audiologie et orthophonie, Université de Montréal, Canada

Abstract

Research on the sensory entrainment of neural oscillations provides a novel way of understanding how the brain processes spoken language without postulates of interim linguistic units. Several reports have shown that oscillations in the theta range (3-10 Hz) are entrained by syllable-size modulations in the energy envelope of speech. This entrainment has recently been shown to provide sensory frames in processing feature-related cues. A similar perspective has been applied to delta waves (< 3 Hz). Thus, it is suggested that delta entrains to long ("sentence"-size) energy contours, and this would provide processing frames involved in utterance comprehension. The present paper adopts a different viewpoint based on our previous work showing that delta-size perceptual chunks in speech relate to an on-line sensory memory of sequential information. We used electro-encephalography (EEG) to monitor listeners (n=18) on-line responses to utterance stimuli with controlled patterns of energy, pitch, and temporal marks. Measures of inter-trial phase coherence in neural oscillations show that delta specifically entrains to temporal chunks in the stimuli. This supports a view that delta is not entrained by sentence-size patterns but by perceptual chunks which provide sensory frames in processing incoming sequential information in heard speech.

Key words: neural oscillations, entrainment, perceptual chunking, sensory memory

Introduction – The role of neural entrainment

Experimentalists who refer to linguistics in elaborating tests with speech stimuli are incessantly confronted to the problem that units of linguistic analyses do not directly reflect in signals. Separate letter-like "phonemes" are not observed in speech acoustics or physiology. It is also acknowledged, that there are no operational definitions of "words" and "sentences" -- except by reference to spaces in text (e.g. Haspelmath, 2011), and *a fortiori* there are no consistent marks for dividing such units across contexts and languages. Finally, to this day, one finds no evidence in neurosciences that supports a processing of utterances in terms of hierarchical combinations of phonemes, words, and sentences, as theorized in formal linguistics (Ingram, 2007). On this problem of directly linking brain processes to speech signals, research on neural entrainment provides a new perspective.

In particular, studies have shown that neural oscillations in the theta range (3-10 Hz) are entrained by syllable-size modulations in the energy envelope of continuous speech. This entrainment, often measured in terms a "cerebro-acoustic coherence", creates theta-length sensory windows which are involved in processing "distinctive features" (e.g. Luo & Poeppel, 2012). To illustrate how this functions, one can refer to temporal indices in speech such as voice onset time (VOT) and transitions, which cue voice and place features (e.g. /ta/-/da/, /ta/-/ka/), respectively. It is known that these temporal indices vary with speech rate in that VOTs and transitions shorten when speaking at fast rates. Yet listeners correctly categorize sounds, basically because the indices are perceived relative to a syllable frame. This is seen in Fig. 1: absolute ms values of VOT vary with speech rate (panel A) but are categorical when one measures VOT as a ratio relative to syllable durations (panel B). In fact, altering the perceptual frame of syllable cycles, without modifying VOT, can influence categorization (Boucher, 2002).

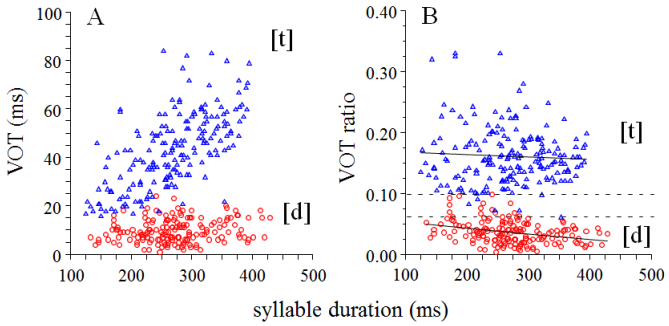


Figure 1. Variations of VOT across speech rates. A-absolute VOTs; B-relative VOTs. A category boundary supporting distinctions is given by B.

A recent brain-imaging study has illustrated this effect in terms of neural entrainment. Thus, ten Oever and Sack (2015) showed that entraining theta oscillations on varying syllable durations can influence the categorization of transitions. In short, indices that cue classical feature distinctions are processed in terms of a *sensory frame* that links directly to theta waves, and there is no need in this approach to postulate interim units of representation like phonemes or groups of consonants and vowels.

As for longer frames, some research suggests that energy envelopes variably related to prosodic groups or "sentences" can provide processing frames that bear on utterance "comprehension" (e.g., Park et al. 2015). In these reports, however, cerebro-acoustic coherence was measured by reference to energy contours of continuous speech. Consequently, the contours could reflect a number of structures in speech. In our approach, we

used utterance stimuli with narrowly controlled pitch, intensity, and temporal patterns so as to identify the effects of given structures on brain responses. We found that an on-line perceptual chunking operates specifically in terms of temporal groups, and presents effects on the sensory memory of heard items in utterances, *whether these contain meaningful or incomprehensible syllables* (Gilbert et al, 2014, 2015). Hence, the perceptual chunking did not relate to utterance comprehension per se, but to a memory processing of incoming sequential information. Moreover, in examining listeners' average brain responses to the stimuli, group-related periodicities appeared that could match oscillations in the delta range – as illustrated in Fig. 2. In order to verify whether such patterns in EEG reflected the entrainment of delta, we performed the present experiment.

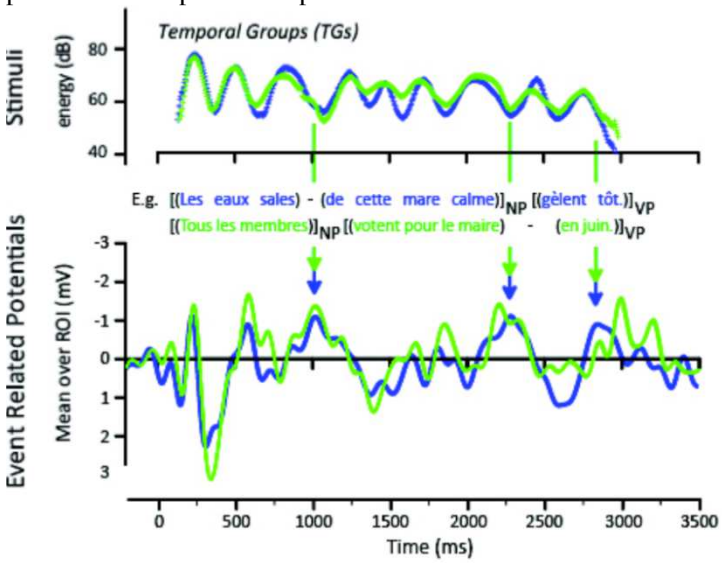


Figure 2. Listeners' ERPs (*n*=20) at electrode Cz follow temporal groups in utterances (50 trials per condition), regardless of whether intonation contours reset on the third (green) or seventh (blue) syllable (from Gilbert et al. 2015)

Method and results -- Delta entrainment

Eighteen listeners heard utterances containing regular temporal groups at periods of 685 ms (1.46 Hz), with narrowly controlled intonation and energy envelopes. Stimuli construction and EEG recording was that of previous studies of Gilbert et al (2014, 2015). In the analyses, however, we calculated the inter-trial phase coherence (ITPC) using EEGLAB (Delorme & Makeig, 2004), which measures how brain oscillations align with the heard stimuli at different frequencies. Fig. 3 presents a time-frequency plot of ITPC. One can see delta-range oscillations at about 1.5 Hz elicited while participants heard

utterances with regular temporal groups of 1.46 ms. Hence the delta-range response was specifically evoked by the periodicities of temporal groups.

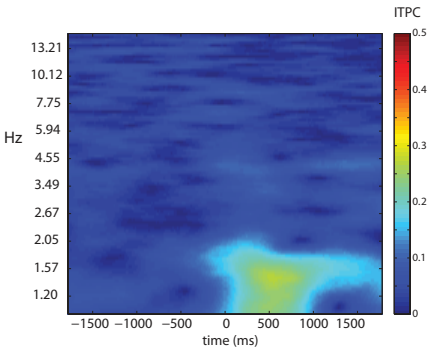


Figure 3. Time-frequency plot of Inter-trial phase coherenc (ITPC) at electrode FC4 while listeners (n=18) heard utterances containing regular temporal groups of 1.46 ms. Peaks in ITPC are in the range of delta oscillations. Utterances were controlled for pitch and presented similar energy enveloppes as in Gilbert et al. (2014, 2015).

Discussion and prospective

The above entrainment was also observed with meaningless series of syllables and, consequently is not involved in sentence comprehension, as suggested by Park et al. (2015). However, delta entrainment was especially clear at fronto-central sites, and this, in conjunction with earlier findings, suggests that delta oscillations may relate to perceptual chunks involved in processing sequential sensorimotor information. Further analyses comparing listeners' responses to speech and non-speech will serve to confirm whether delta entrainments bears specifically on a processing of sensorimotor aspects as opposed to a processing of acoustic information.

References

Boucher, V.J. 2002. Timing relations in speech and the identification of voice-onset times. *Perception & Psychophysics* 64, 121-130.

Gilbert, A.C., Boucher, V.J., Jemel, B. 2015. The perceptual chunking of speech: A demonstration using ERPs. *Brain Research* 1603, 101-113.

Gilbert, A.C., Boucher, V.J., Jemel, B. 2014 Perceptual chunking and its effect on memory in speech processing. *Frontiers in Psychology 5*,1-9.

Haspelmath, M. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45, 31-80.

Ingram, J.C.L. 2007. Neurolinguistics. An introduction to spoken language processing and its disorders. Cambridge UK: Cambridge University Press.

Luo, H., Poeppel, D. 2012 Cortical oscillations in auditory perception and speech... *Frontiers in Psychology* 3,170.

Park, H. et al 2015 Frontal top-down signals increase coupling of auditory low-frequency oscillations. *Current Biology* 25, 1649–1653.

Ten Oever, S., Sack, A.T. 2015 Oscillatory phase shapes syllable perception. *P. N. A. S.* 112, 15833-15837.

Open vs. closed syllable phonology and temporal production in Greek

Anthi Chaida[1,2], Ilectra Dimoula[2], Evgenia Magoula[2], Olga Nikolaenkova[3]
[1]Lab of Phonetics & Computational Linguistics, University of Athens, Greece
[2]Department of Primary Education, University of Athens, Greece
[3]Department of General Linguistics, Saint Petersburg State University, Russia

Abstract

The present experimental study examines syllable constituent durations as a function of syllable structure, lexical stress and focus. In accordance with a production experiment, the results indicate that both syllable structure and lexical stress have significant effects on syllable constituent durations but not focus. Syllable structure has a compensatory temporal effect, according to which open syllables have longer nucleus vowels but shorter onset consonants in comparison to closed syllables. Lexical stress has a lengthening effect on all syllable constituents, in the order nucleus vowel > onset consonant > coda consonant. Onset consonants and nucleus vowels showed significant interactions between syllable type and lexical stress, which indicates extensive variability of segment duration in accordance with prosodic structure and prosodic context of spoken utterances.

Key words: consonant, vowel duration, syllable, temporal production, Greek

Introduction

The present study is an experimental investigation of segment durations as a function of open vs. closed syllable structure as well as lexical stress and focus.

Research on different languages has concentrated on vowel nucleus duration, suggesting that vowels in open syllable context are longer than respective vowels in closed syllable context (Maddieson 1985). The vowel lengthening effect of open syllabicity has even been phonologised in Swedish, where open syllables have phonological long vowels whereas closed syllables have phonological short vowels. This open syllable vowel lengthening effect, either phonological or phonetic, remains an open issue as results from different studies have been fairly controversial (McCrary 2004).

In this study, in addition to the open vs. closed syllable temporal variability of vowel nucleus, considerable attention is paid to each and every syllable constituent, including syllable onset in both open and closed syllable contexts as well as syllable coda in closed syllable contexts. Another aspect, which has hardly drawn particular attention in the international literature, is the effect of variable prosodic conditions, particularly lexical stress and focus, on each syllable constituent.

Experimental methodology

The speech material consists of four test words: two in nominative singular and two in accusative plural with lexical stress at the antepenultimate and penultimate, respectively (Table 1). The test words were produced at the context of the carrier phrase [ˈfonakse ___ ðinaˈta] ‘s/he shouted ___ loudly’. Five female speakers at their early twenties, with standard Athenian Greek pronunciation, produced the speech material at a normal tempo in focus and out of focus contexts in a sound-treated studio at Athens University Phonetics laboratory. The speech material was analysed with the Praat programme and the results of segment duration measurements were subjected to statistical processing with SPSS statistical package.

Table 1. Test words in nominative and accusative with lexical stress assignment in antepenultimate and penultimate syllable, respectively.

Nominative singular	Accusative plural	Gloss
ˈsinɵetos	sinˈɵetus	Synthetic
ˈsisomos	siˈsomus	Whole

Results

The results are shown in figures 1-3. In accordance with a three-way Anova (syllable type x lexical stress x focus), syllable type and lexical stress have significant effects on both onset consonant and vowel nucleus durations whereas focus has no significant effect on any syllable constituent.

Figure 1 shows mean durations of syllable type constituents. Onset consonant in open syllables is 89 ms (SD 20) and in closed syllables 97 ms and this difference of 8 ms is significant ($F=4.5$, $p<0.03$). Nucleus vowel in open syllable is 76 ms (SD 22) and in closed syllable 55 ms and this difference of 22 ms is also significant ($F=30$, $p<0.0001$).

Figure 2 shows mean durations of syllable constituents as a function of lexical stress application. Stressed syllables are 219 ms (SD 30) and unstressed syllables 147 ms, a significant difference of 72 ms ($F=129$, $p<0.0001$). Onset consonants in stressed syllables are 107 ms and in unstressed syllables 79 ms, a significant difference of 28 ms ($F=97$, $p<0.0001$). Vowel nucleus in stressed syllables are 82 ms and in unstressed syllables 47 ms, a significant difference of 35 ms ($F=133$, $p<0.0001$. Coda consonants in stressed syllables are 57 ms and in unstressed syllables 47 ms, a significant difference of 10 ms ($F=11$, $p<0.01$).

Figure 3 shows mean durations of syllable constituents as a function of focus application. Although syllables and syllable constituents in focus are

slightly longer than respective syllables and syllable constituents out of focus, any duration differences do not reach significance level.

The results indicate significant interactions of syllable type x lexical stress with reference to onset consonant ($F=4.2$, $p<0.04$) but not any other syllable constituent.

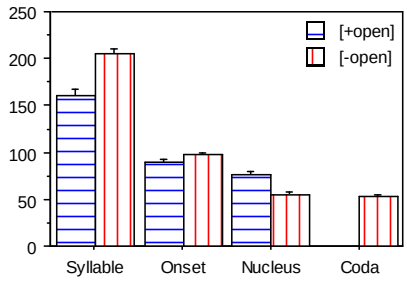


Figure 1. Onset, nucleus and coda syllable constituent durations (in ms) as a function of open (+open) vs. closed (-open) syllable type.

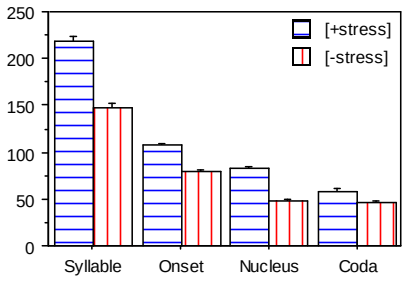


Figure 2. Syllable, onset, nucleus and coda constituent durations (in ms) as a function of stressed (+stress) vs. unstressed (-stress) syllables.

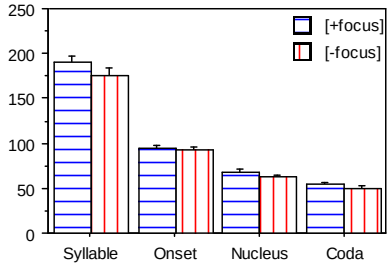


Figure 3. Syllable as well as onset, nucleus and coda constituent durations (in ms) as a function of focus (+focus) vs. out of focus (-focus) context.

Discussion and conclusions

The present study has produced new findings as well as corroboration of earlier results. The new findings are related to (1) the onset consonant and nucleus vowel compensatory duration tendencies, i.e. open and closed syllabicity implies onset consonant and nucleus vowel mirror-image duration pattern and (2) the variability of lexical stress lengthening effects according to syllable structure constituency, i.e. vowel nucleus > onset consonant > coda consonant. On the other hand, the lengthening effect of lexical stress application on onset consonant and vowel nucleus as well as the absence of any lengthening effect of focus application on any syllable constituent have been corroborated (Botinis 1989, Fourakis, Botinis, Katsaiti 1999).

A widely discussed phonetic correlate of syllable structure is duration variability, according to which open and closed syllables show respective tendencies of increased and decreased vowel durations (Maddieson1985). In Swedish, vowel length is claimed to be distinctive, in accordance with a rhyme VC complementary distribution. Thus, short vowels (V) are followed by long (geminate) consonants or consonant clusters (CC) and long vowels (V:) are followed by short (singleton) consonants (C), which results to VC.C vs. V:.C (VV.C) syllable structures, e.g. “vil:a” (villa) “vi:la” (rest). In Italian, on the other hand, consonant length is claimed to be distinctive and a question is thus raised in respect to the preceding vowel length (Bertinetto 2004, Farnetani and Kori 1986). Several studies report on vowel lengthening as a function of open syllabicity, e.g. “ca:ne” (dog), and vowel shortening as a function of close syllabicity, e.g. “car.ne” (meat) (Farnetani and Kori 1986). On the other hand, Farnetani and Kori (1986) report consonant lengthening at coda position, which is in complementary distribution with vowel duration.

Regardless the phonological status of consonants or vowels and their respective duration correlates in Italian and Swedish, Greek has no consonant or vowel length distinctions and any variability of segment durations is phonetic at segment level. On the other hand, as the results of the present study imply, segment durations reflect different syllable constituents, which are phonological rather than phonetic at syllable level.

Acknowledgements

Thanks to Yi Xu for comments and much useful feedback as well as Maria Papanagiotou and Athina Tsiori for data recordings and data analysis. The Special Account for Research Grants, National and Kapodistrian University of Athens, has granted research and travel allowance.

References

Bertinetto, P.M. 2004. On the undecidable syllabification of /sC/ clusters in Italian: Converging experimental evidence. Italian Journal of Linguistics / Rivista di Linguistica, 16, 349-372.

Botinis, A. 1989. Stress and Prosodic Structure in Greek. Lund University Press.

Farnetani, E. and Kori, S. 1986. Effects of syllable and word structure on segmental durations of spoken Italian. Speech Communication 5, 17-34.

Fourakis, M., Botinis, A., Katsaiti, M. 1999. Acoustic characteristics of Greek vowels. Phonetica 56, 28-43.

Maddieson, I. 1985. Phonetic cues to syllabification. In Fromkin, V. (ed.), Phonetic Linguistics, 203-221. London: Academic Press.

McCrary, K.M. 2004. Reassessing the role of the syllable in Italian phonology: An experimental study of consonant cluster syllabification, definite article allomorphy and segment duration. Ph.D. thesis, UCLA.

Surveying Greek language instructors' beliefs about metaphor teaching

Ioannis Galantomos
Department of Mediterranean Studies, University of the Aegean, Greece

Abstract

The goal of this paper is twofold. Firstly to explore Greek language instructors' beliefs about the place of figurative language in foreign language pedagogy and secondly to examine their beliefs about metaphor teaching. It is widely accepted that language instructors bring to the classroom various views about all aspects of their work. In L2 instruction, vocabulary is an important component of a learner's communicative competence. Within vocabulary, metaphor awareness contributes to native-like fluency. Given the pervasiveness of metaphor in ordinary communication, we designed a small-scale study in order to survey Greek instructors' beliefs about its place in Greek as an L2 instruction. Our results are mixed. Pedagogical implications are discussed.

Key words: Greek as an L2, instructors' beliefs, vocabulary, metaphor teaching

Introduction

Over the past two decades, research has focused on the study of L2 teachers' beliefs (Niu & Andrews 2012). Language instructors bring to the classroom certain views about all aspects of their work. These beliefs have to do with knowledge, their students, their discipline and instructional methods (Levin 2015). Teachers' beliefs are not static, but dynamic, context-related and shifting (Johnson 2009).

The major tenor of teacher beliefs' studies is related to the teaching of grammar, reading and writing (Borg 2006) with relatively few studies investigating L2 vocabulary and teachers' beliefs (Rossiter, Abbott & Kushnir 2016). L2 vocabulary is deemed to be an important component of a learner's communicative competence (Read 2004). Within vocabulary, metaphor comprehension and production contribute to native-like fluency (MacLennan 1994).

Given the above, we conducted a small-scale experiment in order to investigate on the one hand Greek language instructors' beliefs about the place of figurative language in L2 instruction and on the other hand their beliefs regarding the teaching of metaphor.

Method

Research questions

Two research questions guided our study:

What are Greek language instructors' beliefs about figurative language (focus on metaphor)?
To what extent are Greek language instructors' beliefs about figurative language consistent with their teaching practices?

Participants

The sample of our study consisted of a total of 13 Greek language instructors who came from private and public Greek language centers in the greater area of Athens and taught at various CEFR (Common European Framework of Reference for Languages) levels. More specifically, 4 of them taught at A2, 3 at B1, 4 at B2 and the remaining 2 at C2 level. Regarding our participants' sex, 5 were males and 8 females. Their mid age was 28.2 years old. In addition, 10 of them were holders of an MA either in Linguistics (Applied or Theoretical) or Education.

Instrument

The metacognitive approach (with a focus on the metacognitive aspect of language learning) was followed to elicit our participants' beliefs. Within this approach, data are gathered through self reports (Barselos 2003).

Procedure

Our participants were asked to express their thoughts in written form with regard to the above mentioned (research) questions. For this assignment they were given 45 minutes. Upon completion, the essays were gathered and their wording was analyzed in order to identify the basic patterns regarding the two research questions.

Results

Our results are mixed. On the one hand, our participants believe that figurative language is closely related to L2 competence (research question 1). On the other hand, almost none of them taught metaphorical meanings and strategies of how Greek learners should deal with unknown figurative vocabulary (Research question 2). In particular, all of our participants (100%) wrote that figurative is an essential component of L2 communicative competence. However, only 4 of them had incorporated figurative language into their syllabus (30%). It is worth mentioning that those 4 instructors

taught at B2 (2 instructors) (15%) and C2 (2 instructors) (15%) levels respectively.

Discussion

Our mixed findings can be attributed to certain beliefs and teaching practices in use in Greece with regard to figurative language. Traditionally, figurative language lies at the periphery of L2 teaching (Thornbury 2002). Nevertheless, recent studies have demonstrated that figurative language is ubiquitous in everyday communication and that it performs key functions in discourse (Semino 2008). Therefore, the ability to use figurative language appropriately can enhance a learners' overall communicative competence (Littlemore & Low 2006). Under this perspective, Littlemore, Krennmayr, Turner & Turner (2014) suggest that learners' familiarization with figurative language should start at CEFR A2 level onwards.

In Greece, the curriculum designed by the Centre for the Greek Language (Gr. *KEI*) (i.e. the official state organization for the world-wide certification for the knowledge of the Greek language) (Αντωνοπούλου, Βογιατζίδου & Τσαγγαλίδης 2013) suggests that instructors should raise their learners' awareness of figurative language at C2 level. In addition and based on empirical data, few textbooks (especially at lower CEFR levels) introduce figurative language to Greek learners. Hence, our findings are not surprising. Our participants' beliefs are in line with those reported in literature (e.g. Kalyuga & Kalyuga 2008) regarding the beneficial contribution of figurative language in L2 vocabulary teaching. On the other hand, present curriculum of Greek as an L2 does not offer learners the opportunity to realize the systematic nature of figurative language.

In light of the above, it is essential for curriculum designers to introduce figurative language from early on, taking into consideration the various functions it serves at different CEFR levels (cf. Littlemore et al. 2014). Additionally, Greek language instructors should raise their learners' awareness of the pervasiveness of figurative language through various activities based on authentic material. This technique has been proven to be beneficial for expanding learners' vocabulary size (e.g. Boers 2004). Finally, specialized seminars should take place so as to familiarize Greek language instructors with the place and functions of figurative language and to revitalize instructional methods.

Conclusions

To sum up, the results of our survey reveal a mixed situation regarding our research questions. On the one hand, it is accepted that figurative language plays an important role in L2 instruction and on the other hand its teaching

has not yet received its proper place in Greek as an L2 instruction. In light of these findings certain teaching implications are discussed.

References

Αντωνοπούλου, Ν., Βογιατζίδου, Σ., Τσαγγαλίδης, Α. 2013. Πιστοποίηση επάρκειας της ελληνομάθειας. Νέο αναλυτικό εξεταστικό πρόγραμμα. Θεσσαλονίκη, ΥΠΑΙΘΠΑ & ΚΕΓ.

Barcelos, A.M.F. 2003. Teachers' and students' beliefs within a Deweyan framework: Conflict and influence. In Kalaja, P., Barcelos, A.M.F. (eds.) 2003, Beliefs about SLA, 7-33. Dordrecht, Kluwer..

Boers, F. (2004). Expanding Learners' Vocabulary Through Metaphoric Awareness: What Expansion, What Learners, What Vocabulary? In Achard, M., & Niemeier, S. (eds) 2004, Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching, 211-232. Berlin & New York, Mouton de Gruyter.

Borg, S. 2006. Teacher cognition and language education: Research and practice. London, Continuum.

Johnson, K.E. 2009. Second Language Teacher Education: A Sociocultural Perspective. New York & London, Routledge.

Kalyuga, M. & Kalyuga, S. 2008. Metaphor awareness in teaching vocabulary. Language Learning Journal 36/2, 249-257.

Levin, B.B. 2015. The Development of Teachers' Beliefs. In Fives, H. & Gill, M.G. (eds.) 2015, International Handbook of Research on Teachers' Beliefs, 48-65. New York & London, Routledge.

Littlemore, J. & Low, G. 2006. Figurative thinking and foreign language learning. Basingstoke, Palgrave Macmillan.

Littlemoe, J., Krennmayr, T., Turner, J. & Turner, S. 2014. An Investigation into Metaphor Use at Different Levels of Second Language Writing. Applied Linguistics 35/2, 117-144.

MacLennan, C.H. 1994. Metaphors and prototypes in the learning teaching of grammar and vocabulary. IRAL 32/2, 97-110.

Niu, R. & Andrews, S. 2012. Commonalities and discrepancies in L2 teachers' beliefs and practices about vocabulary pedagogy: A small culture perspective. TESOL Journal, 6, 134-154.

Read, J. 2004. Research in Teaching Vocabulary. Annual Review of Applied Linguistics 24, 146-161.

Rossiter, M.J., Abbott, M.L. & Kushnir, A. 2016. L2 Vocabulary Research and Instructional Practices: Where Are the Gaps? TESL-EJ 20/1, 1-25.

Semino, E. 2008. Metaphor in Discource. Cambridge, CUP.

Thornbury, S. 2002. How to teach vocabulary. Harlow, Pearson Longman.

Producing appropriate prosodic cues in a non-dominant language: preliminary results from French-English bilinguals

Annie C. Gilbert[1,2], Max Wolpert[1,3], Shanna Kousaie[2,4], Shari R. Baum[1,2]
[1] School of Communication Sciences and Disorders, McGill University, Canada.
[2] Centre for Research on Brain, Language, and Music, Canada.
[3] Integrated Program in Neuroscience, McGill University, Canada.
[4] Montreal Neurological Institute and Hospital. McGill University, Canada.

Abstract
Adapting one's production of prosodic cues to a second or non-dominant language can be difficult. The present study focuses on French-English bilinguals' ability to adapt their prosody to coordinate phrase-final lengthening and lexical stress. Because French has no lexically-coded prosody, it might be difficult for French-dominant speakers to simultaneously control lexical and phrasal prosodic cues. Our preliminary results demonstrate that not only the speaker's L1, but the relative dominance of one language over another can predict speakers' ability to adapt prosody to the specific demands of different languages, at least with respect to controlling syllable duration. These findings are in line with recent results showing that native French listeners do not process lexical stress automatically, instead relying on alternative perceptual mechanisms.

Key words: Bilingualism, L2 prosody, stress, lengthening, language dominance.

Object of study
Adapting one's production of prosodic cues to a second or non-dominant language can be difficult and may lead to a perceived foreign accent or even render speech hard to interpret. The present study focuses on French-English bilingual speakers' (FEs) ability to adapt their prosody to coordinate phrase-final lengthening and lexical stress. Because French has no lexically-coded prosody, it might be difficult for French-dominant speakers to simultaneously control lexical and phrasal prosodic cues. The goal of our study is to examine how FEs with different patterns of language dominance use duration to mark lexical and phrasal prosody in English and French.

Methodology
Participants
Nineteen healthy FEs were recruited from the Montréal area. Information about participants' first and second languages (L1 and L2) and language learning history was gathered through questionnaires, and an index of language dominance was obtained by comparing the English and French

scores on an adapted sentence recall subtest of the Clinical Evaluation of Language Fundamentals 4 (Semel, et al., 2003).

Stimuli

The task involved sentence pairs containing the same two-syllable string presented either as two monosyllabic words separated by a phrasal boundary (e.g. English: *If you want a **key, we** can duplicate one*. French: *Le vendeur d'**or loge** à l'hôtel*.) or as one bisyllabic word preceding a phrasal boundary (e.g. English: *If you want a **kiwi**, I can buy one tomorrow.* French: *Le vendeur d'**horloge** vit à l'hôtel*.). Thirty sentence pairs were created in each language (English and French).

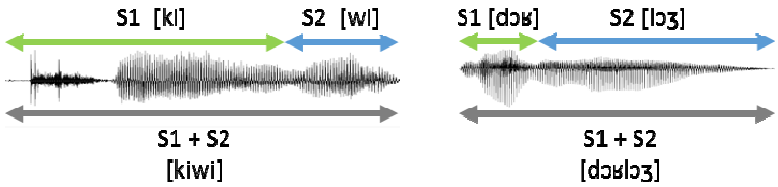
Procedures

Sentences (pseudorandomized but blocked by language) were presented on a computer screen, and participants were instructed to read them aloud in a neutral tone. Trials with disfluencies, pauses, or production errors were discarded, along with the matching sentence in the pair.

Analysis

To control for speech rate variability, relative duration measures were computed by comparing the duration of each syllable of the two-syllable string (S1 e.g.: [ki] and S2 e.g.: [wi]) to the duration of the entire ambiguous region (S1+S2, see Figure 1 for a schematic representation of the measurements). Using these values, we calculated an asymmetry score by subtracting the relative duration of S1 from the relative duration of S2 (henceforth "S2-S1 asymmetry"). A positive S2-S1 asymmetry score indicates that the second syllable was relatively longer than the first syllable, and vice versa. All statistical analyses were performed using *Generalized Linear Mixed Effects* models (Baayen, et al., 2008; implemented in the *lme4* package (Bates, et al., 2014), version 1.1-7, in *R* (R Dev. CoreTeam, 2010), version 3.1.3).

Significance levels were obtained using the *lmerTest* package (Kuznetsova, et al., 2015), version 2.0-29. This method was chosen over traditional statistical analyses because it takes into account subject-related variability, which represents a great advantage when working with small numbers of participants. Models were fitted with *subjects* and *sentence pairs* as random effects (with random intercepts) and *language produced* (English or French), *number of words in the ambiguous region* (one bisyllabic word – i.e. *kiwi* –, or two monosyllabic words – i.e. *key, we*) and participant's language dominance index as fixed factors.



English example French example

Figure 1. Two examples of S2-S1 asymmetry calculation. Left panel represents an English example with a negative S2-S1 asymmetry score, where S2 is shorter than S1, whereas the right panel represents a French example with a positive S2-S1 asymmetry, where S2 is longer than S1.

Results

Statistical models revealed a significant three-way interaction among target language (English vs French), number of words in the ambiguous region (1 bisyllabic vs 2 monosyllabic words), and participants' language dominance index in predicting S2-S1 asymmetries [β < -0.07112, *SE* = 0.025, t = -2.824, p < 0.01] (see Fig. 2). Interestingly, the pattern of S2-S1 asymmetries remained stable across speakers in French trials, but was modulated by participants' language dominance index for English trials. English-dominant and non-English-dominant speakers (balanced bilinguals or French-dominant) behaved differently from the English-dominant speakers when speaking in English. In contrast to English-dominant speakers, non-English-dominant speakers showed a smaller S2-S1 asymmetry within bisyllabic word trials, such that S1 and S2 were approximately equal in duration.

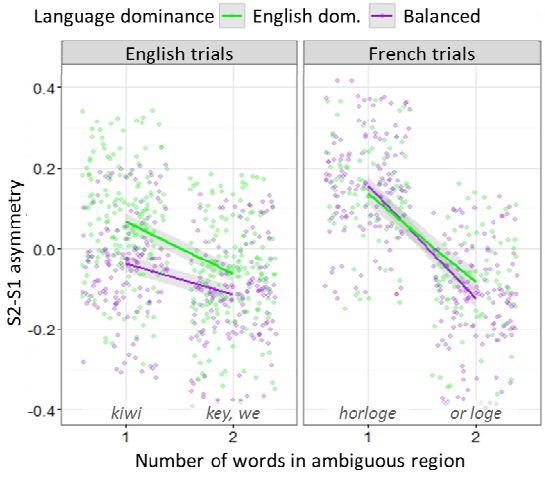


Figure 2. Interaction between the number of words in the ambiguous region (x axis), language of the trial (English – left panel, French – right panel) and estimated language dominance (English dominant in green, balanced bilinguals in purple). Fitted data points and linear regressions from the LME model.

Similar results were obtained if the language dominance index is replaced by the speaker's self-reported L1 [β < -0.1252, *SE* = 0.057, t = -2.198, p < 0.03]. That is, only English L1 speakers pattern with English dominant speakers and manage to modulate the duration of the two syllables in the bisyllabic word condition. Non-native English speakers and/or non-English dominant speakers seem to have a difficult time differentiating the duration changes associated with lexical stress from the duration changes associated with phrase final lengthening.

Conclusions

The preliminary results demonstrate that not only the speaker's L1, but the relative dominance of one language over another can predict speakers' ability to adapt prosody to the specific demands of different languages, at least with respect to controlling syllable duration. These findings are in line with results reported by Michelas et al. (2016), showing that, unlike native English listeners, native French listeners do not process lexical stress automatically, instead relying on alternative perceptual mechanisms.

Acknowledgements

Funding for this research was provided by the Fonds de Recherche du Québec - Société et Culture, The Blema and Arnold Steinberg Family Foundation, and The Centre for Research on Brain, Language and Music.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. 2008. Mixed-effects modeling with crossed random effects for subjects and items. J. Mem. Lang., 59(4), 390-412.

Bates, D., Maechler, M., Bolker, B., & Walker, S. 2014. lme4: Linear mixed-effects models using Eigen and S4. (Version R package version 1.1-6). Retrieved from http://CRAN.R-project.org/package=lme4.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. 2015. lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-29.

Michelas, A., Frauenfelder, U. H., Schön, D., & Dufour, S. (2016). How deaf are French speakers to stress? JASA, 139(3), 1333-1342.

R Development Core Team. 2010. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from www.R-project.org

Semel, E., Wiig, E. H., & Secord, W. A. 2003. Clinical Evaluation of Language Fundamentals 4 (CELF-4). . San Antonio, TX.: The Psychological Corporation.

Production of voice onset time (VOT) in Austrian German conversational speech: a pilot study

Petra Hödl
Department of Linguistics, University of Graz, Austria

Abstract

This paper deals with voice onset time (VOT) in Austrian German conversational speech. It presents methods and results from a pilot production study designed to investigate which VOT values native speakers of Austrian German typically produce for realisations of /p, t, k/ and /b, d, g/. Preliminary results from one female talker show a lot of variability in VOT values for voiceless plosives. Furthermore, for the bilabials and alveolars there is a considerable amount of overlap between phonologically voiced and voiceless plosives. This might suggest a tendency towards merging these two categories in conversational speech.

Key words: acoustic phonetics, voice onset time, Austrian German, conversational speech.

Introduction

Standard German exhibits a two-way phonemic contrast between so-called "voiced" (/b, d, g/) and "voiceless" (/p, t, k/) stops. However, similar to English, in word-initial position this phonemic contrast is usually realised as an aspiration contrast rather than a "true" voicing contrast. Hence, phonetically we deal with different durations of positive voice onset time (VOT) – the time interval between the stop release and the beginning of quasi-periodicity in the speech signal (Lisker & Abramson 1964). In doing so, we distinguish between voiceless "lenis" stops (with a short-lag VOT) and voiceless aspirated "fortis" plosives (with a long-lag VOT).

This is certainly true for the pronunciation of stops by North German speakers. However, for speakers of some (Southern) varieties of German, such as Austrian German, the situation might not be so clear. According to the literature, these varieties may exhibit a reduced degree of aspiration or even show a total lack of it (Muhr 2001: 798-799, Krech et al. 2009: 239).

In contrast to this claim, previous production studies found statistically significant differences between VOTs for word-initial fortis and lenis plosives in Standard Austrian German (Grassegger 1988, Moosmüller & Ringen 2004). While these studies dealt with read speech, the aim of the current paper is to investigate VOT in conversational speech. The main goal is to figure out whether in Austrian German aspiration of voiceless stops is restricted to careful speech or whether it occurs to the same degree in conversational speech.

Methods

To investigate VOT in conversational speech, a production study was designed which involved an interactive picture-describing task.

The design of the task was based on the diapix technique (e.g. Baker & Hazan 2011), which was developed to elicit dialogues between two talkers. During the task, each talker is given one of two almost identical pictures containing a certain number of differences. Since the talkers do not see the picture of their partners, they have to verbally describe them in order to spot as many differences as possible. In the current design, one talker (A) is assigned the role of the main describer, while the other talker's (B) task is to find and mark the differences.

Material

For the proposed study, four DIN A4 picture-duplets with various scenes were designed. These scenes contained drawings of 166 target words, i.e. German words with initial voiced or voiceless plosives followed by either the vowel /a/, /i/ or /o/. 97 items carried word-initial stress, the rest did not. It has to be noted that the number of words in each category was not balanced. This limitation of the design was due to the restriction to words that could be drawn as well as due to gaps in the lexicon.

Subjects

In order to be able to elicit conversational speech, pairs of talkers are recorded together. Talker A, who is the person whose recordings are analysed, is asked to bring a friend or colleague along to participate in the picture-describing task. All subjects are native speakers of Austrian German with self-reported normal hearing and no speech or language disorders.

Here, data from pilot recordings done with one female talker (F01), aged 56, are presented. This talker grew up in Graz, the second biggest city in Austria, located in the South-East of the country. She speaks with an accent close to standard Austrian German with hardly any dialectal features.

Recordings

Recordings were made under living-room conditions in a quiet room at the Department of Linguistics at the University of Graz. Two lavalier microphones and an external audio-interface were used to conduct two-channel recordings at a sampling frequency of 44.1 kHz and a quantization level of 16 bit.

Acoustic analysis

For data analysis the produced target words were manually segmented and annotated in Praat (Boersma & Weenink 2017). VOT measurements were taken from the onset of the burst to the first indication of voicing in the speech signal. For judging the position of the burst and the beginning of voicing the waveform as well as the spectrogram of the signal were used.

Preliminary results

From the 166 target words talker F01 produced 138 words. Some of them were uttered more than once, which yielded the total number of 183 analysed items for this talker. For general analysis, the three vowel contexts and the two stress contexts were collapsed together.

As can be seen in Figure 1, VOTs for word-initial plosives show a fair amount of variability. Additionally, one can see the tendency for velar plosives to have longer VOTs than bilabials or alveolars. This is in accordance with the well-documented effect of place of articulation on aspiration and duration of VOT (Cho & Ladefoged 1999).

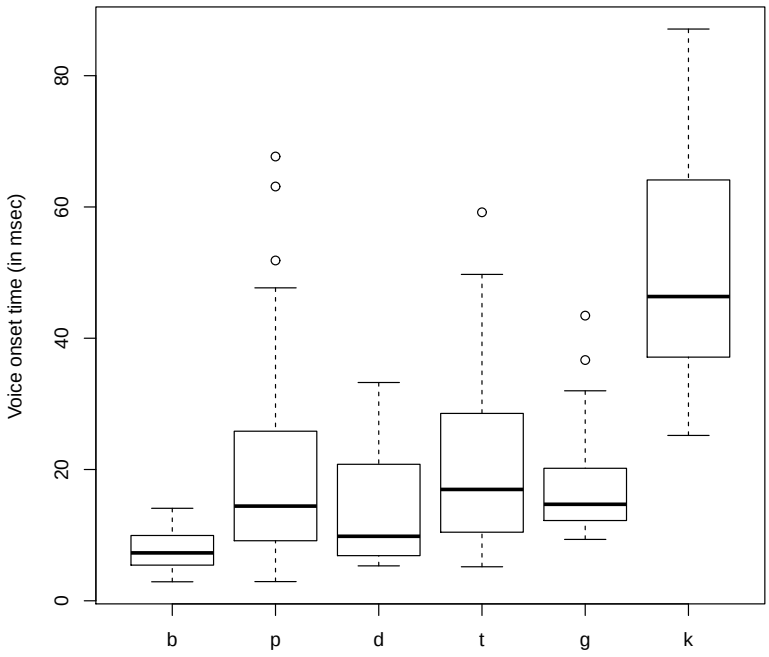


Figure 1. Boxplots of VOT values of speaker F01 for realisations of /b/, /p/, /d/, /t/, /g/ and /k/. The dark horizontal line depicts the median, the boxes the interquartile range (IQR) and the whiskers the minima and maxima (within 1.5 x IQR). The circles depict outliers outside of 1.5 x IQR.

Another observation was that – while realisations of /b, d, g/ and /p, t. k/ both showed positive VOT values (hence, voicing lag) – voiceless plosives tended to exhibit longer positive VOTs than their voiced counterparts. This is especially true for velars (see Figure 1).

However, for the bilabial and alveolar plosives there seems to be a certain degree of overlap between the two voicing categories. While there were items in which initial /p/ and /t/ were produced with strongly positive VOTs, there were also many items that were clearly not.

Conclusion

Since data collection and analysis are currently still under their way, definitive answers to the research question cannot yet been given. What is striking, though, is the considerable amount of variance of VOT values for /p, t, k/ observed in this pilot data.

Crucially, for bilabials and alveolars there even seems to be a certain degree of occasional overlap between VOTs for "voiced" and "voiceless" plosives. While there is aspiration of a number of word-initial plosives in these pilot recordings, there is a lot of variation concerning the degree of aspiration. This might indeed suggest a tendency towards occasionally merging these two categories in Austrian German conversational speech.

References

Baker, R., Hazan, V. 2011. DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. Behav. Res. Methods 43, 761-770.

Boersma, P., Weenink, D. 2017. Praat: doing phonetics by computer [computer program]. Version 6.0.28, http://www.praat.org/.

Cho, T., Ladefoged P. 1999. Variations and universals in VOT: evidence from 18 languages. Journal of Phonetics 27, 207-229.

Grassegger, H. 1988. Signalphonetische Untersuchungen zur Differenzierung italienischer Plosive durch österreichische Sprecher (= Forum Phoneticum 40). Hamburg, Buske.

Krech, E.-M., Stock, E., Hirschfeld, U., Anders, L. C. (eds.) 2009. Deutsches Aussprachewörterbuch. Berlin/New York, de Gruyter.

Lisker, L., Abramson, A. S. 1964. A cross-language study of voicing in initial stops: acoustical measurements. Word 20, 384-422.

Moosmüller, S., Ringen, C. 2004. Voice and aspiration in Austrian German plosives. Folia Linguistica 38, 43-62.

Muhr, R. 2001. Varietäten des Österreichischen Deutsch. Revue belge de philologie et d'histoire 79, 779-803.

Posh accent and vocal attractiveness in British English

Li Jiao[1], Chengxia Wang[2], Cristiane Hsu[2], Peter Birkholz[3], Yi Xu[2]
[1]School of Foreign Languages, Tongji University, China
[2]Dept of Speech, Hearing and Phonetic Sciences, University College London, UK
[3]Inst of Acoustics and Speech Communication, Technische Universität Dresden, Germany

Abstract

Posh accent in British English is associated with upper class. Previous research on poshness has been centred on vocabulary, grammar and phonology, but little is known about the phonetic properties. This study, as part of a larger project, is an attempt to connect posh accent with attractiveness of voice through a common set of dimensions originating from emotional prosody research. Using VocalTractLab and Praat, we created stimuli varying in voice quality, nasality, formant shift ratio, pitch shift and duration. Results of two separate perception experiments showed that only voice quality and formant shift ratio functioned significantly. Breathy voice sounded the most posh and attractive, and pressed voice the least. Likewise, utterances with the smallest formant shift ratio sounded the most posh and attractive.

Key words: poshness, attractiveness, voice quality, formant shift ratio

Introduction

Poshness, or posh accent, is a British notion that a person sounds upper class when speaking. Previous studies on poshness have centred on vocabulary, phonology and grammar (Ross 1954, 1970, Fox 2004). One is heard as posh if s/he speaks clearly and uses correct grammar and certain words. However, there have also been anecdotal claims about phonetic properties associated with a posh accent that have not yet been systematically studied.

One language coach suggests that speaking further forward in the mouth and holding tight the lips make a person sound posh (YouTube 2015). Another suggests speaking as slowly as possible (YouTube 2013). Yet a third claims that posh people hold back their emotions in public (YouTube 2014), which implies that they also sound cold and detached. Finally, there are also claims that link nasality to posh accent (Herbert 1988, Buuren 1988).

In this study, which is part of a larger project (Jiao *et al*. in press), we explore these ideas with a method developed for studying affective prosody and vocal attractiveness (Xu *et al*. 2013). The method is based on the hypothesis that emotions and attractiveness are both vocally encoded with acoustic parameter that project a large or small body size to dominate or appease the listener (Ohala 1984).

Method

Stimuli

A short statement *You are feeling mellow*, read by a 16-year-old Cambridge-accented male speaker from the Intonational Variation in English (IViE) corpus (Grabe *et al*. 1998), was used as a template for articulatory synthesis using VocalTractLab (Birkholz 2013). Six base sentences were created that varied in voice quality (breathy, modal and pressed) and nasality (-nasality and +nasality). In the next step, these six base sentences were modified with the PSOLA algorithm implemented in Praat (Boersma 2001) to generate a total of 162 stimuli that differed as follows:

Formant shift ratios ranging from compressed to expanded spectrum (0.9, 1.0, 1.1), stimulating a lengthened or shortened vocal tract.
Pitch shifts ranging from lowered to raised (-3, 0, 3) median pitch of the utterance in semitones.
Duration ratios ranging from shortened to lengthened (0.8, 1.0, 1.2) total duration of the utterance.

Participants

Seventeen native English speakers (average age: 24.2; 8 females), who self-reported to be able to tell a posh accent, took part in the experiments in a lab at the Department of Speech, Hearing and Phonetic Sciences, UCL and a quiet study room in Scape Shoreditch, London.

Perception tasks

Two experiments were run in which subjects judged how posh (experiment 1) and how attractive (experiment 2) each utterance sounded on a 5-point scale. There were 324 trials in each experiment as all stimuli were heard twice in random order. Each utterance was heard only once in each trial.

Results

Repeated Measures ANOVAs showed main effects of voice quality and formant shift ratio for both poshness [voice quality ($F(2, 32) = 7.848$, $p = 0.0017$); formant shift ratio ($F(2, 32) = 6.356$, $p = 0.0047$)] and attractiveness [voice quality ($F(2, 32) = 21.899$, $p < 0.0001$); formant shift ratio ($F(2, 32) = 9.896$, $p = 0.0005$)]. Student-Newman-Keuls tests showed that breathy voice received significantly the highest scores for both poshness and attractiveness. For the other voice qualities, pressed was rated lower than modal in both tasks although the difference in poshness was rather small. Likewise, utterances with the largest formant shift ratio (1.1) sounded the least posh and attractive. Although there was no significant difference

between the smaller ratio (0.9) and the original (1.0), their average scores showed a favour for the longer vocal tract (0.9) in both tasks. There were also marginal four-way interactions among voice quality, nasality, formant shift ratio and duration for both poshness ($F(8, 128) = 2.056$, $p = 0.0449$) and attractiveness ($F(8, 128) = 2.566$, $p = 0.0125$), which are too complex and subtle to interpret.

Apart from the above, there was also a marginal interaction between voice quality and nasality ($F(2, 32) = 3.642$, $p = 0.0376$) for poshness. For attractiveness, there was a main effect of duration ($F(2, 32) = 3.67$, $p = 0.0367$), whereby the shorter durations (0.8 and 1.0) had significantly higher scores, and the 0.8 ratio had the highest mean score. Additionally, for attractiveness, there were two-way interactions between nasality and pitch shift ($F(2, 32) = 5.646$, $p = 0.0079$), formant shift ratio and pitch shift ($F(4, 64) = 11.73$, $p < 0.0001$) and formant shift ratio and duration ($F(4, 64) = 2.964$, $p = 0.0261$) as can be seen in Figure 1. Due to lack of space, however, these two-way interactions can not be further discussed.

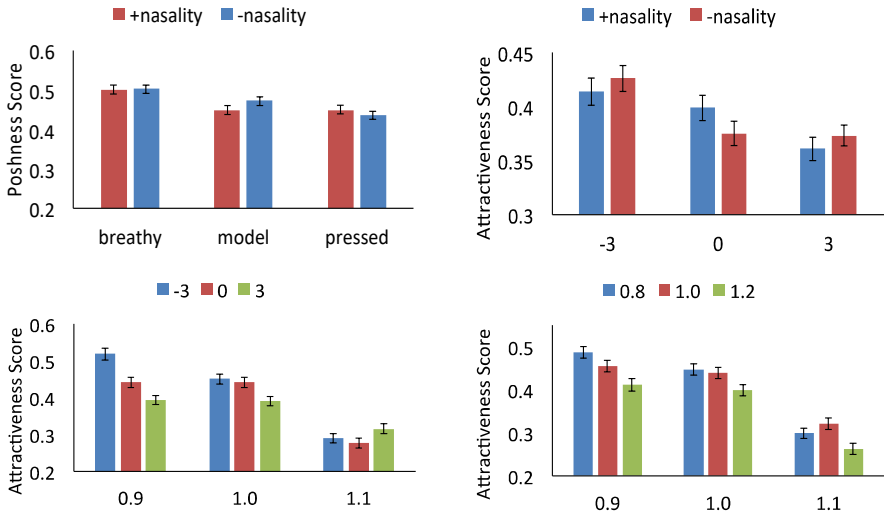


Figure 1. Upper left: Normalized poshness scores as a function of voice quality & nasality; Rest of the plots: Normalized attractiveness scores as a function of nasality & pitch shift (upper right), formant shift ration & pitch shift (lower left) and formant shift ratio & duration (lower right).

Discussion and conclusion

The results of the two experiments are partially consistent with the impressionistic accounts of posh accent in terms of social connotations (but not phonetic properties). That is, with a lengthened vocal tract that projects a

large body size, it sounds dominant and authoritative, qualities that happen to be also associated with an attractive male voice (Xu *et al*. 2013). On the other hand, both poshness and attractiveness are associated with a breathy voice, which was found in attractive male as well as female voices (Xu *et al*. 2013). Posh accent is therefore kind of like an attractive male voice, masculine but not aggressive. Contrary to the anecdotes, nasality has no main effects on either poshness or attractiveness, although it is involved in some interactions, which can be further explored in future studies.

Acknowledgements

Thank Mia and Judy whose posh English play inspired this project.

References

Birkholz P. 2013. Modeling consonant-vowel coarticulation for articulatory speech synthesis. PLoS ONE, 8(4), e60603.

Boersma, P. 2001. Praat, a system for doing phonetics by computer. Glot International 5, 341-345.

Buuren, L. Van 1988. Margaret Thatcher's pronunciation: an exercise in ear-training. Journal of the international phonetic association 18 (01), 26-38.

Fox, K. 2004. Watching the English. London: Hodder and Stoughton Ltd.

Grabe, E., Nolan, F. and Farrar, K. 1998. IViE - a Comparative transcription system for intonational variation in English. In Mannell, R. and Robert-Ribes, J. (ed.), Proceedings of 5th International Conference on Spoken Language Processing, 1259-1262, Sydney, Australia.

Herbert, J. 1988. Broadcast speech and the effect of voice quality on the listener: a study of the various components which categorise listener perception by vocal characteristics. PhD thesis, University of Sheffield.

Jiao, L., Wang, C., Hsu, C., Birkholz, P. and Xu, Y. in press. Does posh English sound attractive? in Proceedings of INTERSPEECH 2017, Stockholm, Sweden.

Ohala, J. 1984. An ethological perspective on common cross-language utilization of F_0 of Voice. Phonetica 41,1-16.

Ross, A. 1954. Linguistic class-indicators in present-day English. Neuphilologische Mitteilungen 55 (1), 20-56.

Ross, A. 1970. How to pronounce it. London: Hamilton.

Xu, Y., Lee, A., Wu, W., Liu, X. and Birkholz, P. 2013. Human vocal attractiveness as signaled by body size projection. PloS ONE, 8 (4), e62397.

YouTube. 2013. How to Speak POSH English. [online] Available at: https://www.youtube.com/watch?v=BLyPhCR4S1M. [Accessed 19 May 2017].

YouTube. 2014. Why Posh People are Cold. [online] Available at: https://www.youtube.com/watch?v=rKAGTKAwgqE. [Accessed 19 May 2017].

YouTube. 2015. How to sound posh - Part one. [online] Available at: https://www.youtube.com/watch?v=Nj0Rh__1kDw [Accessed 19 May 2017].

Acquisition of L2 English adverb placement by L1 Russian and Greek Cypriot speakers

Sviatlana Karpava
University of Central Lancashire, Cyprus

Abstract

This study investigates the acquisition of L2 English adverb placement by L1 Russian and L1 Greek Cypriot speakers. It aims to find out whether structural differences and similarities among English, Russian and Greek lead to facilitative or negative syntactic transfer effect in L2 acquisition of English. The results of the grammaticality judgment task showed that there is a facilitative transfer from L1 Russian into L2 English and negative transfer from L1 CG into L2 English. It was found that the participants had a better performance on the adverbs of frequency rather than on adverbs of manner. L2 proficiency and years of L2 learning affect adverb placement and the rate of cross-linguistic influence.

Key words: adverb placement, word order, L2 English, Russian, Cypriot Greek.

Introduction

This study investigates the acquisition of L2 English adverb placement by L1 Russian and L1 Greek Cypriot (CG) speakers. It aims to find out whether structural differences and similarities among English, Russian and Greek lead to facilitative or negative syntactic transfer effect in L2 acquisition of English.

According to Alexiadou and Anagnostopoulou (1998), the difference between Greek and English can be explained by universally fixed adverb positions and verb movement. Greek is a null subject language, with V-movement. Verb-adverb word order is licensed in Greek, but not in English or Russian, while adverb-verb word order is allowed in Russian and English (Kallestinova and Slabakova, 2008; Mykhaylyk et al., 2015), but not in Greek or CG.

The unmarked position of frequency adverbs in Greek is post-verbal (SVadvO), while pre-subject, sentence initial (AdvSVO) is marked, but not completely impossible. The grammatical position of manner adverbs is post-object (SVOadv), whereas pre-subject position (AdvSVO) is marked (Alexiadou and Stavrakaki, 2005).

In English, the unmarked position of frequency adverbs is pre-verbal (SadvVO), though post-object and pre-subject are allowed as well, while the grammatical position of adverbs of manner is post-object (SVOadv), though pre-verbal and pre-subject placement is also licensed. Post-verbal position,

both for adverbs of manner and frequency, in the sentences with direct object is ruled out.

The Generalized Inversion Hypothesis (Bailyn, 2004) suggests that in Russian SVO sentences adverbs of frequency and manner are allowed in pre-verbal position but not in post-verbal, though the latter is not considered fully ungrammatical and has degraded status (Kallestinova and Slabakova, 2008).

This study is an attempt to look into the complex process of L2 English acquisition of adverb placement by L1 Russian and CG speakers, whether there is cross-linguistic interference on syntax-semantics interface and whether adverb (mis)placement depends on the type of adverb, VP-internal and VP-external, age, length of exposure to L2 and level of L2 proficiency.

Methodology

The participants of the study were 100 CG undergraduate students, 17-24 years old, 50 males and 50 females. The other group of the participants were 25 L1 Russian learners of L2 English, 17-28 years old, 19 females and 6 males. The L2 English proficiency of the participants was measured by IELTS test: listening, reading, writing and speaking.

The grammaticality judgement task was implemented, which had the following conditions: (i) pre-subject position of adverb, sentence initial position, allowed in English, Greek and Russian, for both manner and frequency adverbs; (ii) pre-verbal position of adverb, grammatical in English and Russian, for both manner and frequency adverbs, and ungrammatical in Greek; (iii) post-verbal position of adverb, ungrammatical in English and Russian, for both manner and frequency adverbs, but unmarked in Greek only for frequency adverbs; (iv) post-object position of adverb, sentence final position, allowed in English and Russian, for both manner and frequency adverbs, but licensed only for adverbs of manner in Greek.

The participants had to rate the grammaticality of the sentences, using Likert scale from 1 (totally incorrect) to 5 (totally correct). The task was focused on adverb placement, but there were also distractor items on the use of tenses and subject-auxilliary inversion. The participants were tested on the placement of adverbs of manner and frequency. The linguistic (socio-economic) background questionnaires were used as well.

Results and discussion

The analysis of the data showed that for the first condition, pre-subject position of adverb, there is no difference between Russian and CG students. Sentence initial position of adverb is allowed in Russian and Greek. Both groups of the participants had higher rating (mean) for the sentences with

adverbs of frequency than adverbs of manner. This could suggest that the type of the adverb might affect the grammaticality judgment of L2 English sentences.

Regarding the second condition, it was found that Russian students scored higher (mean rating scores of the grammaticality judgment task) in comparison to CG students. This could be explained by positive transfer from L1 Russian and negative transfer from L1 CG. Also, the rating for the adverbs of frequency were higher than for the adverbs of manner, for both groups of the participants.

There was not revealed any difference between Russian and CG students with respect to the third condition, post-verbal position of adverb. Both L1 Russian and L1 CG learners of L2 English scored lower for the adverbs of frequency than for the adverbs of manner.

Concerning the fourth condition, Russian students scored higher for the sentences with adverbs of manner and lower than CG students for the sentences with adverbs of frequency. The performance of the participants on this condition could be explained by positive transfer from L1 Russian and L1 CG. The details are provided in Table 1.

Table 1. Grammaticality judgment task, adverb placement, mean scores.

Adverb position	Adverbs	Russian students	CG students
Pre-subject	Manner	2.67	2.65
	Frequency	3.17	3.15
Pre-verbal	Manner	3.95	3.27
	Frequency	4.41	3.68
Post-verbal	Manner	3.35	3.28
	Frequency	3.03	3.01
Post-object	Manner	4.07	3.67
	Frequency	3.18	3.37

According to one-way ANOVA, years of exposure to L2 English, years of learning L2 English, is a statistically significant factor for pre-verbal placement of frequency adverbs *(F(24)= 34.778, p=.007)* in the grammaticality judgment task by Russian students.

IELTS scores, level of L2 English proficiency, is a statistically significant factor for pre-subject placement of manner adverbs *(F(99)= 3.666, p=.002)* and pre-subject placement of frequency adverbs *(F(99)= 3.141, p=.007)* in the grammaticality judgment task by CG students.

According to paired samples t-test, there is a statistically significant difference between manner and frequency adverbs in pre-subject position *(t(99)= -5.663, p=.000)*; pre-verbal position *(t(99)= -2.730, p=.009)*; post-

verbal position *(t(99)= 2.577, p=.014)* and post-object position *(t(99)= 2.708, p=.010)* in the grammaticality judgment task by CG students; in pre-subject position *(t(24)= -2.592, p=.032)*; pre-verbal position *(t(24)= -3.496, p=.008)* and post-object position *(t(24)= 3.286, p=.011)* in the grammaticality judgment task by Russian students.

The results showed that with respect to the grammaticality judgment task there is a facilitative transfer from L1 Russian into L2 English and negative transfer from L1 CG into L2 English. Overall, Russian students performed better than CG students.

It was found that L2 learners of English had a better performance with respect to the adverbs of frequency than the adverbs of manner, which could be the evidence that verb movement is sensitive to adverb hierarchy: frequency>manner (Cinque, 1999).

L2 proficiency, length of exposure to L2 English, years of L2 learning, affect adverb placement and the rate of cross-linguistic influence. L1 transfer is more evident at the early rather than at the advanced stages of L2 acquisition.

References

Alexiadou, A. and Anagnostopoulou, E. 1998. Parametrizing Agr: word order, verb-movement and EPP-checking. Natural Language and Linguistic Theory 16, 491-539.

Alexiadou, A. and Stavroula Stavrakaki, S. 2005. Functional categories and clause structure in a Greek-English speaking bilingual patient with Broca's Aphasia: Evidence from adverb placement. In Cohen, J., McAlister, K., Kellie Rolstad, K. and MacSwan, J. (eds.), ISB4: Proceedings of the 4th International Symposium on Bilingualism, 39-57. Somerville, MA: Cascadilla Press.

Bailyn, J. 2004. Generalized inversion. Natural Language and Linguistic Theory 22, 1-49.

Cinque, G. 1999. Adverbs and Functional Heads: A Cross-linguistic Perspective. New York: Oxford University Press.

Kallestinova, E. and Slabakova, R. 2008. Does the verb move in Russian? In Antonenko, A., Bailyn, J. and Bethin, C. (eds.), Proceedings of FASL 16, 199-214. Michigan Slavic Publications.

Mykhaylyk, R., Mitrofanova, N., Rodina, Y. and Westergaard, M. 2015. The Linguistic Proximity Model: The case of Verb-Second revisited. In Grillo, E. and Jepson, K. (eds.), Proceedings of the 39th annual Boston University Conference on Language Development, 337-349, Sommerville, MA, Cascadilla Press.

Substantial factors for decision-making in syllabification in Russian

Galina Kedrova, Valentina Kolybasova
Lomonosov Moscow State University, Russian Federation

Abstract

Present study investigates significant factors essential for decision-making in syllabification in Russian. Words representing various types of actual phonotactic constraints and regular phonetic processes in Russian speech has been elected as experimental dataset. Results of syllabification experiments reveal preferred individual syllabification strategies, among the most stable should be mentioned tendency for predominance of V-coda syllables over C-codas. Other influential principles of syllabification are: 1) predomination of consonantal clusters in syllable's onset, 2) tendency to avoid syllables consisting of a single vowel, 3) a kind of a glottal stop at the beginning of vowel's articulation at the word's onset. Orthographic and phonetic transparency/opacity of a word has also been stated one of the crucial factors for syllables boundaries in consonantal clusters.

Key words: Syllabification; Russian language, grapheme/phoneme misalignments.

Introduction

The role of syllable as core element in speech production has been universally acknowledged in the phonetic science. However, results of multiple syllabification experiments using either straightforward or very sophisticated experimental procedures and techniques are rather ambiguous. Therefore, despite all efforts to explain mechanism(s) and principles of syllabification in various languages productive and/or perceptive role of a syllable as fundamental linguistic notion is still unclear and generally admitted to raise intricate problems (Krakow 1999). One of the most challenging is the problem of verification of syllabification principles in various languages (Goslin 2001).

Whereas at the very beginning of experimental research activity in the field most theories of syllable and syllabification procedures were based mainly on the English language data, in the past two decades emerged more experimental evidence for various world languages. Most of them support idea of syllable as language-dependent special linguistic segmentation unit as well as minimal motor planning component of speech articulation. Russian language data on syllables typology and syllabification behaviour in Russian speech, which dates from early 1990s, is abundant and diversified (Bondarko 2000). However, main stumbling block of all concepts of the Russian syllabification strategies deals with so-called 'syllabification

boundaries problem'. While every Russian speaker unambiguously identifies number of syllables in a Russian word (with very few and very well known exceptions), she/he regularly disaccords on indicating syllables boundary within a word. Thus, following numerous experimental data Zlatoustova put forward a proposal on fundamental impossibility to match syllable boundaries in the Russian.

To distinguish languages of similar type from syllable-counting languages, Kodzasov proposed a special term for the languages with intrinsic 'floating' syllable boundaries: a 'wave' language (unlike 'quantum' language). While most experimental data evidenced in favour of prevailing type of the open syllable (CV) model in Russian, many linguists yet consider idea of universality of the CV-model for the Russian language still a challenging one. Therefore, the main purpose of the present study was to discern potential influence of phonetic processes in consonant and vocalic clusters in Russian speech on syllabification strategies of speakers.

Experimental data and procedures

Our research differs from previous ones in approach to experimental stimuli phonetic parameters. Experimental stimuli dataset composition was composed of Russian words representing various types of phonotactic constraints and phonetic transformations occurring in Russian speech (vowel reduction, sonorisation / devocalization, regular assimilation patterns occurring in consonantal clusters, dissimilation, hiatus, consonantal dropouts and schwa insertion). The dataset was balanced according to degree of orthographic and phonetic transparency/opacity of a word (mismatch between spelling and pronunciation) while it has been already demonstrated that factor of transparent *vs* opaque orthography proved its validity for experimental subjects' behaviours in French (Chetail 2012).

The experimental set of word items comprised 653 words with 1.893 potential syllables equal to the number of vowel characters (1.903 syllables considering syllables with verified schwa insertions). The dataset was sequentially reorganized according to the principle of accrescent pronunciation and reading complexity. All word items were split into 3 word sets, each presented to experimental subject in consequent sub-series during one experimental session. Each series word set also differed on proportion of size and constitution of consonant clusters (Table 1).

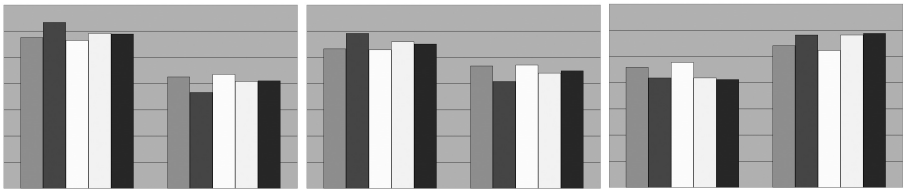
Table 1: Experimental stimuli data parameters. Legend for data in Table 1: N(w) – number of words; N(s) – number of syllables; N2 – 2-component consonantal clusters; N3 – 3-component consonantal clusters; N4 – 4-component consonantal clusters; N5 – 5-component consonantal clusters; N(a) – phonetic transformations in consonant clusters.

Series	N(w)	N(s)	N2	N3	N4	N5	N(a)
1	218	216	626	15	0	0	30
2	219	650	708	27	8	1	57
3	70	350	569	53	12	0	107

Subjects included 5 adult native speakers of Russian (4 males and 1 female) of ages between 31 and 65 with various professional and educational experiences. Three of the experimental subjects were professional linguists (2 of them phoneticians); one subject was mathematician and one a physicist. A list of words split into three quasi-equal sub-lists (series) was presented to experimental subjects on a computer screen to be syllabified orally. After completing the main experimental task on syllabification subjects were asked to read experimental stimuli from the screen as naturally as possible and at own pace. All the speaking activities has been recorded and later transcribed manually by the professional phoneticians.

Results

None of the experimental subjects posed any questions concerning experimental instruction, while the very idea of ‘syllable’ as word constituent has been generally introduced in Russian elementary school education. However, main results of segmentation of test words into syllables by experimental subjects revealed considerable discrepancies in individual syllabification strategies and syllables’ inventory for every speaker. Results of segmentation of experimental words into syllables according to syllable codas (V/C) for every participant in all experimental sessions are presented in Figure 1.



Session 1 Session 2 Session 3

Figure 1. V-codas (left cluster) and C-codas syllables (right cluster) for every subject in each session.

Conclusions

Analysis of the results of recorded and transcribed experimental subjects syllabification behaviours reveals personal preferences for V-coda syllables *vs* C-coda syllable. Other influential principles of syllabification are: predomination of consonantal clusters in syllable's onset *vs* syllable's codas; tendency to avoid syllables consisting of a single vowel or syllables with vowel onset; glottal stop at the beginning of vowel's articulation at the word's onset. These findings corroborate previous reports on maximum onset principle as a universal one for structurally different languages. Hypothesis of pre-planning strategies underlying execution of oral syllabification task depending on language writing system (Chetail 2006) was also strongly supported by our results. Russian words without or with minimum orthographic mismatch tend to be segmented in a more uniform manner, while dataset with less opaque spelling caused considerable inconsistency within a subject's individual syllabification behaviours alongside with more homogeneous syllabification decisions in more complicated cases for all subjects.

References

Bondarko, K.B., Skrelin, L.V., Shalonova, P.A. 2000. Statisticheskie kharakteristiki otkrytyh slogov. VOX. Budapest, 2000.

Chetail, F., Content, A. 2012. Segmentation of Written Words in French. Language and Speech 56, 125-144.

Chetail, F., Scaltritti, M., Conten, A. 2006. Effect of the consonant-vowel structure of written words in Italian. Quarterly Journal of Experimental Psychology 67 (5), 833-42.

Goslin, J., Frauenfelder, U.H. 2001. A comparison of theoretical and human syllabification. Language and Speech 44, 409-436.

Krakow, R.A. 1999. Physiological organization of syllables: a review. Journal of Phonetics, 27, 23-54.

Affricates and affricated consonants in Aromanian spontaneous speech

Anastasia Kharlamova
Department of General Linguistics, Saint Petersburg State University, Russia

Abstract

This paper deals with affricates and affricated consonants in Aromanian, an Eastern Romance language and member of the Balkan Sprachbund. The materials for the study were obtained in Greece, FYRM, and Albania. The current goal is to analyze the processes that concern affricates in Aromanian spontaneous speech, the main of which are the loss of stop phase in dental affricates and the affrication of dental stops before front-row vowels. Furthermore, we hope to determine the reflection of the speakers' bilingualism in the phonetics of their speech.

Keywords: phonetics, affricates, dental consonants, spontaneous speech, Balkan languages

Introduction

Aromanian spontaneous speech has practically never been the subject of a separate study before, the only known exceptions being our own research from previous years that have dealt with the topic (Харламова 2015; Kharlamova 2016). Spontaneous speech studies have mostly been conducted using the data of more widespread languages such as English, German, Russian, or French.

This paper is dedicated to affricates and affricated consonants in the spontaneous speech of Aromanians in Turia (Greece), Resen (FYRM), and Elbasan (Albania). The Turia materials were gathered in 2002 by an expedition supported by Kleiner Balkansprachatlas project (see Bara et al. 2005 for detailed description); the data from Resen and Elbasan was obtained during our own expeditions in 2015 and 2016 respectively.

The goal at our current stage of research is to fully analyze the phonetic processes that affricates and affricated consonants undergo in Aromanian spontaneous speech.

Affricates in Aromanian

Phonology of Aromanian is a very underdeveloped field of study, with practically all research done in that area concerning only separate dialects, many of which are vastly different from each other in matters such as pronunciation of diphthongs, central vowels etc. It is usually acknowledged that there are four affricates in Aromanian: [ts], [dz], [tʃ], and [dʒ] (Нарумов 2001). However, our analysis of spontaneous speech also provided

occurrences of [t’s’], [d’z’], [cç], and [ɟʝ] (the latter two are most probably a result of Albanian influence since they occur across Albanian dialects).

The two main processes that involve the affricate inventory in Aromanian spontaneous speech are the loss of stop in dental affricates and the affrication of dental stops before front-row vowels.

Having transcribed our texts using Sound Forge and Speech Analyzer, we now seek to analyse these processes, mainly the factors that cause them.

Phonetic factors

The loss of stop in affricates is quite widespread typologically (Kümmel 2007: 376–409), to the point that it is often believed to be a regular stage of affricate evolution. In our materials we have already seen, during our previous stages of research, that the stop in dental affricates is usually lost in short and frequently used words such as *ț́i* ‘what’. Now our purpose is to find out how it depends on phonetic surroundings and on the lexeme’s origin.

Currently we can say that affricates in Romance words experience the loss of stop more often than affricates in borrowings. The main reason for it, as we believe, is the fact that most of the frequent words containing dental affricates are originally Romance. As for the phonetic surroundings, we should say that an affricate is more likely to lose its stop: a) before a central vowel (such as *ḑāsi* ‘said’ pronounced as [zəs’]) b) after [n] (such as *munț́i* ‘mountains’ pronounced as [muns]) c) in the end of the word (multiple examples of verb forms with an orthographic *ț́* in the end were pronounced in our data with [s]).

As for the affrication of [t], it is limited and occurs only before front-row vowels ([i], [e], [ea], [y]). The evidence from our data suggests that a preceding hissing consonant might also serve as a factor of the affrication (such as [ʃ] in *aromanešti* ‘Aromanian’, pronounced as [arumənеʃt’s’i]).

We are going to further analyze the phonetic conditions of the loss of stop and affrication, taking into consideration suprasegmental factors such as word stress and the consonant’s position in relation to it.

Extralinguistic factors

In addition to the phonetic factors listed above, we would also like to mention some extralinguistic factors that, as we believe, may influence the pronunciation of affricates in different lexemes.

First, rare affricates [tʃ] and [dʒ] mostly occur in loanwords, many of which have Turkic origin (or get borrowed from Turkic languages via Albanian). Numerous Turkic borrowings (for example, *xhami* ‘mosque’), first, are pronounced in practically the same way in all Balkan languages, second, particularly in Aromanian, tend to belong to a limited semantic field

(Islamic terminology, some terms of handicraft etc.). Therefore, these words don't belong to the active vocabulary of most of our informants (Aromanians are Orthodox Christians and therefore may avoid Turkic words for religious reasons as well), and they are not involved in phonetic processes of spontaneous speech as they are usually pronounced "carefully", for example, in response to a linguist's phonetic questionnaire.

Second, we should also notice the status of the Aromanian language. It is officially a national minority language only in FYRM; however, cultural activity of Aromanians in Albania has also risen in recent years. In Greece, many Aromanians identify themselves as the Greeks and their language as Greek (more information on the problem of the Aromanians' national identity can be found in Nedelkov 2009).

Our informants in Resen, FYRM, are the only family in their town to use the language, and even they hardly use it spontaneously and mostly preserve it as heritage. From our Albanian informants only one (an elderly man) speaks Aromanian fluently; the next generation (his two nieces) only remembers some words. Only in Turia could truly bilingual Aromanians be found, speaking both Aromanian and Greek from their childhood (Bara et al. 2005).

Therefore, a big part of our data consists of spontaneously read texts rather than "truly spontaneous" speech. Although it has been noted that spontaneous reading has the necessary level of spontaneity to be included in spontaneous speech studies (Богданова-Бегларян и др. 2013: 150), with a language like Aromanian which is not often in common use we should remember that phonetics of spontaneous reading is less likely to have active reduction and sound change processes that are characteristic for the "truly spontaneous speech".

There is also the fact that Aromanian texts are usually written in a "literary" form[1] which is intended to be as purely Aromanian as possible, while spoken Aromanian is the speech of bilinguals and always experiences the results of close language contact. The problem of the Aromanians' bilingualism is the subject of our future steps of research, as well as the question of whether a bilingual possesses a single phonetic system or a separate one for each language.

Notes

1. There is no officially acknowledge literary or standard Aromanian.

Acknowledgements

This paper was completed as part of a research project supported by RSF 14-18-01405 grant. I would like to thank my informants, the Andonovski family

from Resen and the Kristo family from Elbasan, whose friendliness, hospitality and cooperation made this study possible, as well as my colleague A.L. Makarova who helped me with finding informants in Resen and my tutor Prof. A.N. Sobolev who provided me with Aromanian audio materials from Turia.

References

Bara M., Kahl T., Sobolev A.N., 2005. Die südaromunische Mundart von Turia (Pindos). München, Biblion Verlag.

Kharlamova A.V., 2016. Affricates in the spontaneous speech of Aromanians in Turia. In: ExLing 2016: Proceedings of 7th Tutorial and Research Workshop on Experimental Linguistics, 67–70. University of Athens.

Kümmel M. 2007. Konsonantenwandel: Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktion. Wiesbaden, Reichert Verlag.

Nedelkov J.L. The Ethnic Code of the Vlachs at the Balkans // EthnoAnthropoZoom. №6 / Ed. A. Svetieva. – Skopje: St. Cyril and Methodius University, 2009.

Богданова-Бегларян Н.В., Бродт И.С., Куканова В.В., Павлова О.В., Сапунова Е.М., Суббота И.А., Филиппова Н.С., Хан Н.А., Чуйко В.М. и др., 2013. Звуковой корпус как материал для анализа русской речи. Часть 1. Чтение. Пересказ. Описание. Санкт-Петербург, Филологический факультет СПбГУ.

Нарумов Б.П., 2001. Арумыский язык/диалект. In.: Языки мира. Романские языки, 636–656. Москва, Academia.

Харламова А.В., 2015. Опыт фонети.ческого анализа арумынской спонтанной речи (на материале южноарумынского говора села Турья). In: XVIII Международная конференция студентов-филологов. Тезисы докладов, 214–215. Санкт-Петербург, Филологический факультет СПбГУ.

The prosodic structure in Basque

Mélanie Lancien[1], Philippe Martin[2],
[1]Laboratoire de Phonétique et Phonologie, UMR 7018, CNRS/Sorbonne Nouvelle, France
[2]LLF, UMR 7110, CNRS / Université Paris Diderot-Paris 7, France

Abstract

Basque is a lexically stressed language where patterns in accented phrases have been reported to change from region to region (Olano, 2006). In order to better understand 1) the positioning of lexical stress in accent phrases and 2) the prosodic characteristics of stressed syllables in the indication of prosodic structures, we recorded 4 Basque native speakers form Alava, Gipuzkoa, and Labourd on a selection of sentences readings. Results showed that data on Gipuzkoa Basque were consistent with the contrast of melodic slope model (Martin, 1975). The lexical stress was found to be on the last syllable of the full lexical item, and not just on its stem. In a case study, we also evidenced a possible prosodic grammar of contours for Gipuzkoa Basque.

Key words: Basque, prosodic structure, melodic slope

Introduction

Basque has been described in detail since the 17th century but mostly in the syntactical and lexical domains. Phonetics studies of the different varieties of Basque started in the late 20's with Gavel (1921) and Larrasquet (1928). These studies were far away from nowadays' experimental phonetics and phonology, and mostly aimed at describing how phonemes were pronounced in different dialects. All studies made on Basque dialects reported a great variation, sometimes even between two very geographically close regions of the Basque country.

Most studies on prosody or accentuation in Basque languages did use the ToBI annotation system, pooling together phonetical and phonological prosodic cues and using a very restrictive set of indexes for pitch height and phrases boundaries. Already in 1729, Larramendi in his grammar of Gipuzkoa Basque, evidenced a prosodic system mainly instantiated by a stress located on the final syllable of the stem (e.g. no stress on the suffix). Hualde (2004) confirmed that this pitch accent was coherent with modern varieties of Gipuzkoa Basque.

To globally summarize the varieties of accentual patterns found in Basque dialects, Michelena (1972, 1977, 2013) described 4 categories of Basque: type 1 mostly spoken on the western side of the Basque country, have a tonal pattern consisting of several high pitch syllables followed by several low pitch syllables forming tonal plateau. Types 2, 3 and 4 have a primary accent

on one syllable per stem with no possible accent clash. These stress accent types are mostly found in the East side of the Basque Country.

Methods

To investigate Basque accentuation and prosody in several dialects, we designed sentences of increasing complexity (such as: Katua jaten du - The cat is eating - → Katu zuri beltza nere amonaren lorategian xaguak jaten ari da. - The small black and white cat is eating mice in my grandmother's garden.) , so that by reading speakers would produce preferably prosodic structures congruent to these sentence syntactic structures. As speakers were all bilingual French/Basque and mostly trilingual (Basque/French/Spanish) we gave them the sentences in French and asked them to translate it into their native variety of Basque.

Our speakers were two females from Alava (Spain) and Labourd (France) and two males from Gipuzkoa (Spain). Recordings were made in a soundproof room with headset cardioid microphones. We then proceeded to a syllabic annotation of each sentence. Fundamental frequency curves were annotated with WinPitch's computer-assisted annotation function. To ensure better reliability in the annotation process, a spectrogram narrow band is simultaneously displayed with the fundamental frequency curve using exactly the same frequency scale.

Results on stress' place

The examination of syllables' duration and pitch showed that lexical words have their last syllable stressed. By last syllable we mean the very last, including the casual morpheme, and not just the last syllable of the word stem (as opposed to Michelena, 1997).

At the utterance level this examination also showed an expansion of duration and a progressive decrease of F0 on the three last syllables of the sentence when the said sentence was an assertion. Thus, when the verb is final, the first syllable of its stem is accented with an increase of duration and an increase of pitch, and its following syllables are lengthened and show a progressive decrease of pitch.

The final slope C0 starts on the last syllable of the verb's stem and spread itself to the last syllable of the verb's. For sentences where there is a continuation rise this rise happens on the two lasts syllable of the lexical word except if the word is disyllabic, in that case the rise is on the last syllable.

When there is a question mark the very last syllable of the sentence has a rising slope. For some sentences where speakers displayed evidence or doubt we also observed a hat slope on the last syllable of the sentence. For the

sentences with focus, the focalization was realized as a hat slope on two syllables: a wide rise on the penultimate syllable and a falling slope on the last.

A tentative prosodic grammar

To go beyond phonetic observations, we attempted to sketch a limited prosodic grammar based on observations made on one single speaker native from Gipuzkoa (Spain).

The basic assumption is that the sentence prosodic structure organizes accent phrases in a hierarchy which is instanced by the melodic contours located on stressed syllables. Instead of ToBI tonal targets, we used the following melodic contours to annotate stressed syllables.

C0 terminal conclusive contour (assigned from perception of the prosodic end of the sentence).
Cc complex non-terminal contour, flat or slightly falling on the stressed vowel, and rising above the glissando threshold on the final syllable, combining both melodic movements if the final syllable is stressed.
C1 rising non-terminal contour, above the glissando threshold (Rossi, 1971).
C2 falling non-terminal contour, above the glissando threshold.
Cn neutralized rising, flat or falling contour, below the glissando threshold.

Contrast of melodic slope, the case of assertions.

For assertions, the C0 on the last syllable is instanced by a falling slope, a complex rising contour (Cc) is displayed on the subject phrase, and the object phrase is displayed with a falling contour (C2).

Contrast of melodic slope, the case of interrogative sentences.

For interrogative sentences, we observed a pattern of contours in which the C0 (conclusive final) is rising, the subject has a falling contour (C2) and the whole subject phrase has a complex rising contour (Cc). When there is a subject and an object, the subject phrase has a falling contour (C2) and the object phrase is produced with a complex rising slope.

Discussion

Our results are consistent with Larramendi's findings, that claimed that the stress was located on the last syllable of words. Beyond stress position, the contours and prosodic patterns we evidenced addressed a new matter: we assume that Cc we observed on subject phrases could be grounded in the influence of Spanish intonation more than Basque prosodic patterns (which we also observed on two female speaker). As most Basque speaker are

native bilingual or trilingual the influence of other languages is a tricky path to investigate.

We also observed that the hypothesis of assumed congruence between the prosodic and syntactic structures may not always be valid (especially when the sentence becomes more complex). Inside every object phrase we observed that contours were either C2 or Cn: for instance, in the phrase “katu zuri beltza” [tu] in “katu” is displayed with a C2 as well as [tza] in “beltza”. We also noticed that kind of repetition in contours for subject phrases (where C2 are replaced by C1), however this did not appear on every subject phrase. In longer phrases, many contours were neutralized, giving form to plateau between the first and the last stressed syllable of the phrase.

Acknowledgements

We would like to thank Lorea, Maitane, Euken and Ibon for their time, patience, and their enthusiasm during the recordings, as well as the Pariseko Euskal Etxea.

References

Gavel, H. 1921. *Eléments de phonétique basque*. Sociedad de Estudios Vascos, Eusko Ikaskuntza

Goldsmith, J.A. 1976. *Autosegmental phonology.* Bloomington: Indiana University Linguistics Club, Vol. 159.

Larrasquet, J. 1928. *Action de l'accent dans l'évolution des consonnes étudiée dans le Basque Souletin: étude expérimentale précédée de recherches expérimentales sur l'état actuel et l'évolution des vélaires dans le même dialecte.*

Larramendi, M.D. 1729. *El impossible vencido*.

Hualde, J. I. 2004. *Basque phonology*. Routledge.

Martin, Ph. 1975. Analyse phonologique de la phrase française. *Linguistics*, 146, 35-68.

Martin, Ph. 2015. The Structure of Spoken Language. Intonation in Romance. Cambridge University Press.

Michelena, Luis. 1972. Etimología y transformación . *Homenaje a Antonio Tovar*, 305-317.

Michelena, Luis. 1977. *La lengua vasca*. Leopoldo Zugaza, Editor.

Michelena, L. 2013. A note on Old Labourdin accentuation. *Anuario del Seminario de Filología Vasca" Julio de Urquijo", 6*, 110-120.

Rossi, M. 1971. «Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole . *Phonetica*, 23(1), 1-33.

WinPitch 2017. ww.winpitch.com.

Comparing syntactic development in adolescents' written texts

Maria Liakou, Georgia Andreou, Fotini Anastassiou
Department of Special Education, University of Thessaly, Greece/ Hellenic Open University, Greece

Abstract

The aim of this study was to investigate whether adolescents reach the linguistic level corresponding to their age and to compare their syntactic abilities in written narrative and non narrative texts. The results showed that generally older adolescents had better syntactic abilities compared to smaller ones. However, an unexpected linguistic development was observed in the smaller age group and in non narrative texts, showing that the production of complex texts may lead to unexpected linguistic levels.

Key words: syntactic development, written texts, adolescence

Introduction

Linguistic processes and syntactic skills continue to grow over the adolescents' years. The central focus of our study is adolescents with "typical language development", and this term refers to the language someone learns under normal circumstances during the period of eleven or twelve and early adulthood (Andreou, Liakou, Galantomos 2017; Beers, Nagy 2009, 2011; Nippold 2007; Paul 2001). During this period, adolescents learn to read and write narrative and non narrative texts among others. Narrative texts are those that reflect mostly past experiences and humans' actions. Non narrative texts focus on the development of arguments and information without relying on narration (Georgakopoulou, Goutsos 1999). An important marker of later syntactic development is clausal density (or subordination index), which increases during the adolescent years (Nippold, Hesketh, Duthie, Mansfield, 2005). Clausal density is defined as "the average number of clauses (main and subordinate) per T-unit" and it is measured by summing the total number of clauses (independent and subordinate) and dividing by the total number of T-units produced in a language sample (Scott 1988). The samples are broken into T-units, defined as an independent clause with any accompanying subordinate (dependent) clauses (Hunt 1970).

Purpose of the study

Based on the above, the aim of this study was to compare adolescents' ability to produce a number of T-units and clauses in relation to their age,

examining the differences between the two genres. Given the difficulty of non narrative texts, it was also predicted that both groups would use more clauses in narrative than in non narrative texts.

Method

Our research took place in public, urban high schools in Greece. Our sample consisted of two groups of 300 typically developing adolescents (N=300). The average age of the first group was 14.5 (N= 150 early adolescents) and of the second group was 16.6 (N=150 late adolescents).

The students were asked to write two different texts, one narrative and one non narrative on the topic of racism, following the instructions given, which were the same for both groups of high-school students. In total, we analyzed 600 written texts, half of them being narrative and half non-narrative.

The Mann-Whitney U Test was used to detect possible differences in written texts between early and late adolescents and the Wilcoxon Signed-Rank Test was used to compare two sets of scores between narratives and non narratives texts that were written by the same participant. SPSS 15 was used for our statistical analysis and Monte Carlo simulation methods were used to obtain the p-value. When the p-value is less than the significance level a (a =0.05), the result is said to be statistically significant.

Results

The statistical analysis of Mann-Whitney (see tables 1, 2) revealed that there is a statistically significant difference in non-narrative texts between the two groups of adolescents in the: a) total clauses of texts and b) clausal density.

Table 1. Mean percentage of text T-units, text clauses and clausal density by age group (Early vs Late adolescents) in narrative texts.

Type of Adolescents	Total T-units of Text	Total Clauses of Text	Clausal Density
Early adolescents	7.4	22.0	3.2
Late adolescents	7.4	21.8	3.1
p-value	0.772 (NS)	0.689 (NS)	0.984 (NS)

Table 2. Mean percentage of text T-units, text clauses and clausal density by age group (Early vs Late adolescents) in non-narrative texts.

Type of Adolescents	Total T-units of Text	Total Clauses of Text	Clausal density
Early adolescents	7.3	21.8	3.2
Late adolescents	7.3	18.6	2.6
p-value	0.637 (NS)	p= 0.001 (S)	p= 0.001 (S)

The statistical analysis of Wilcoxon (see Figures 1 & 2) showed a statistically significant difference between the two genres in the a) total clauses of text of late adolescents and b) clausal density of late adolescents.

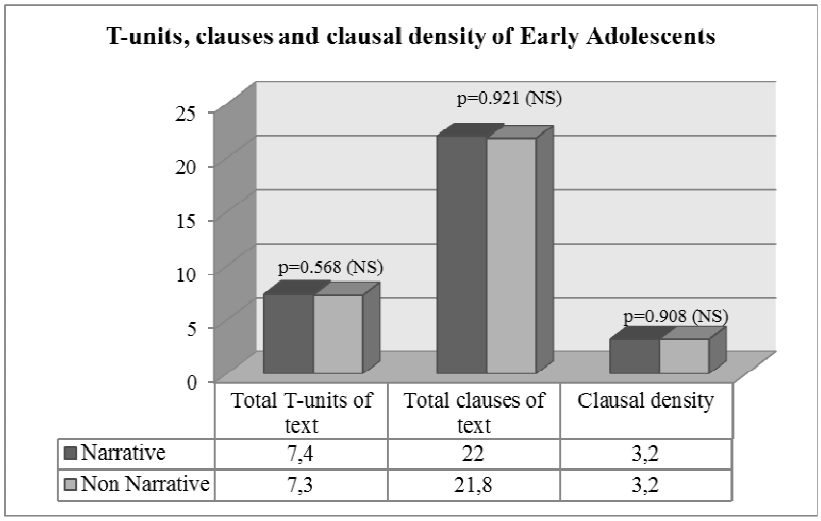


Figure 1. Mean proportion of T-units of text, total clauses of text and clausal density in early adolescents' texts (S: significant, NS: non significant).

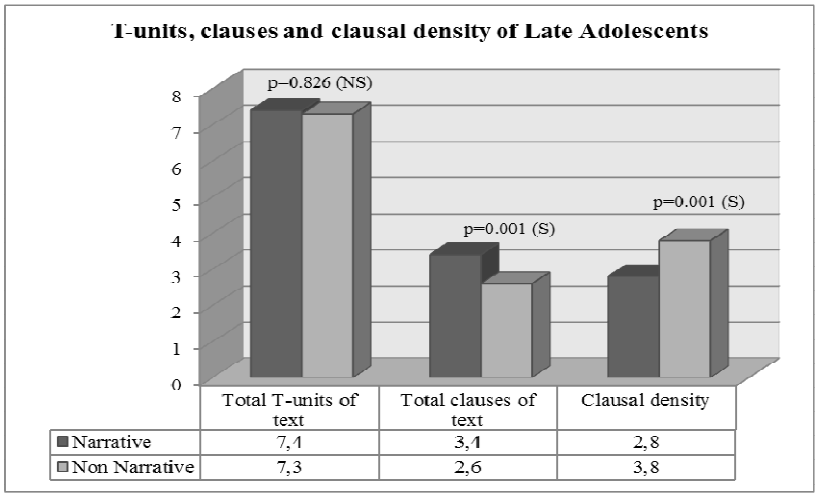


Figure 2. Mean proportion of T-units of text, total clauses of text and clausal density in late adolescents' texts (S: significant, NS: non significant).

Conclusions – Discussion

The results showed great syntactic complexity in narrative and non narrative texts for all age groups. Early adolescents functioned at higher levels of linguistic development than late adolescents regarding all of the criteria. So, our results showed that some students functioned at lower levels of language development corresponding to their age, whereas others reached an advanced linguistic level with considerable syntactical skills. In addition, an unexpectedly high syntactic development of the adolescents in our sample was observed, both in terms of age and the two types of texts examined.

So, syntactic complexity increases during adolescence, marked especially by a growth in the use of clauses (Andreou, Liakou, & Galantomos, 2017). Generally, earlier studies revealed a complex syntax in the areas of narrative and non-narrative texts as well (Beers & Nagy, 2009, 2011; Nippold, 2007; Nippold et.al., 2005).

References

Andreou, G., Liakou, M. & Galantomos, I. (2017). Differences in syntactic development in adolescence. European Journal of Language Studies, 4 (2), 6-11.

Beers S. F. & Nagy W. E., (2011). Writing development in four genres from grades three to seven: syntactic complexity and genre differentiation. Reading and Writing, 24, 183-202.

Beers S. F. & Nagy W. E., (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? Reading and Writing, 22, 185-200.

Georgakopoulou, A. & Goutsos, D. 1999. Text and Communication. Athens: Ellinika Grammata.

Hunt, K. W. (1970). Syntactic maturity in school children and adults. Monographs of the Society for Research in Child Development, 35 (Serial No. 134, No. 1), 1-67.

Nippold, M. 2007. Later Language Development: School-Age Children, Adolescents, and Young Adults. Austin, TX, Pro-Ed.

Nippold M.A., Hesketh L.J., Duthie J.K. & Mansfield T.C., (2005a). Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. Journal of Speech, Language, and Hearing Research, 48, 1048-1064.

Paul, R. 2001. Language Disorders from Infancy through Adolescence (Second Edition). United States of America, Mosby.

Scott, C. M. (1988). Spoken and written syntax. In M. A. Nippold (Ed.), Later language development: Ages nine through nineteen (pp. 49-95). Austin, TX: Pro-Ed.

An acoustic study of Mandarin rhotic suffix

Anqi Liu
Department of linguistics, University Paris Diderot (Paris 7), France

Abstract

In this study, I examine the effects of Mandarin rhotic suffix on the quality of the preceding vowels. Results from a comprehensive acoustic study including monophthongs, diphthongs, and nasal codas show that the rhotic suffix centralizes the preceding vowels, and lowers their F3, across the board. In addition, /i/-final diphthongs and /u/-final diphthongs show different effects of the suffix. Also, adding suffix deletes the alveolar nasal coda /n/, but not the velar coda /ŋ/.

Key words: Mandarin, rhotic suffix, acoustics, vowel quality

Introduction

Mandarin rhotic suffix /ɚ/ (儿 "child") is historically a diminutive marker. In contemporary Mandarin, it lost its grammatical function and became a stylistic marker, which is often considered as a feature of Beijing dialect (Duanmu 2007, Chao 1968) or the speech of Mandarin speakers from northern China (Lu 1995). Previous studies on this suffix primarily focus on its phonological behavior (e.g., Duanmu 2007, Tian 2007). According to Duanmu (2007), Mandarin allows up to two segments in rime, and the rhotic suffix adds to the coda position. When the stem already has two rime segments (e.g., nasal coda or diphthong), the last segment of the stem differ from [retroflex] is replaced by the rhotic (e.g., /an/+/r/→/a^{r}/; /ai/+/r/→/a^{r}/). Although the previous studies explain the phonological behavior of this suffix, they pay little attention to the phonetic aspects of the rhotic suffix. This study aims to provide a more comprehensive analysis on the acoustic consequences of the rhotic suffix on the quality of the preceding vowels, including the stem syllables with a nasal coda or a diphthong, to which the phonological analyses draw specific attention. To my knowledge, Huang (2010) is the only study that examines the acoustic consequences of rhotic suffix on the preceding vowels, but it is limited only to monophthongs.

Methods

Participants and stimuli

Participants were 13 Mandarin native speakers (4M/9F, age range: 25-30), 11 of whom were from Shenyang (northern China) and 2 from Hebei (near Beijing). Shenyang speakers were recorded in a quiet room in Shenyang, with Edirol by Roland (R-09HR) recorder. Hebei speakers were recorded in Paris, in a sound-attenuated booth, using a notebook PC, an external

microphone (Audio Technica AE4100), and a preamplifier (Roland Quad-Capture). All participants speak English as the second language. Three of the participants, including the two who were tested in Paris, have learned French in the university. Another two have learned Japanese in the university.

Stimuli were 78 Mandarin words, a half of which had the rhotic suffix while the other half did not. The stimuli included 13 possible rimes of Mandarin /a, i, u, y, ɤ, ai, əi, au, əu, an, aŋ, əŋ, uŋ/ preceded by an onset consonant /t/, with the exception of /y/: /l/ was used before /y/ because Mandarin phonotactics does not allow /ty/. Three tones of Mandarin (T1, T2, and T4) were included in the current study. Tone 3, the low-dipping tone, was excluded because it is usually realized with a creaky voice (e.g., Belotel-Grenié and Grenié 1994), which makes it not suitable for formant analyses.

Participants were asked to produce each stimulus in a carrier phrase "wo214 ba214 __ xie214 hao214" (/wɤ˨˩˦ pa˨˩˦ __ ɕjɤ˨˩˦ xau˨˩˦/ "I write __ well."). The stimuli were repeated five times in a random order.

Measurements

Formants (F1, F2, F3) were measured for each vowel of all stimuli. First, the vowels were segmented from the onset of F1, to the offset of F2 (rimes without the rhotic suffix and diphthongs with the suffix), or to the onset of F3 drop (monophthongs with the suffix). Then, for monophthongs, formants were taken at the temporal midpoint of each vowel (Zee and Lee 2001) by hand. For diphthongs, formants were extracted at every 5% point of the duration of the vowel in Praat (Boersma and Weenink 2017).

Results

Monophthongs

Figure 1 shows the F1 and F2 of the monophthongs with and without the suffix. The vowel space reduced with the rhotic suffix. Statistical results from linear mixed effect models (predictors: suffix condition, tone, gender; by-speaker random intercept) suggest that the rhotic suffix changed the preceding vowel quality in terms of F1, F2, and F3. First, F1 significantly increased with suffix for the non-low vowels /i, u, y, ɤ/ and decreased for the low vowel /a/ (p's < 0.0001). Second, F2 became significantly higher for the back vowels /u, ɤ/ and lower for the front vowels /i, y/ (p's < 0.0001). The effect of suffix on /a/'s F2 interacted with speaker gender: F2 increased for males but decreased for females. This may be due to the gender difference in /a/'s F2 without the suffix (see Figure 1, top and bottom left). Finally, the suffix significantly decreased F3 for all the monophthongs (p = 0.0005).

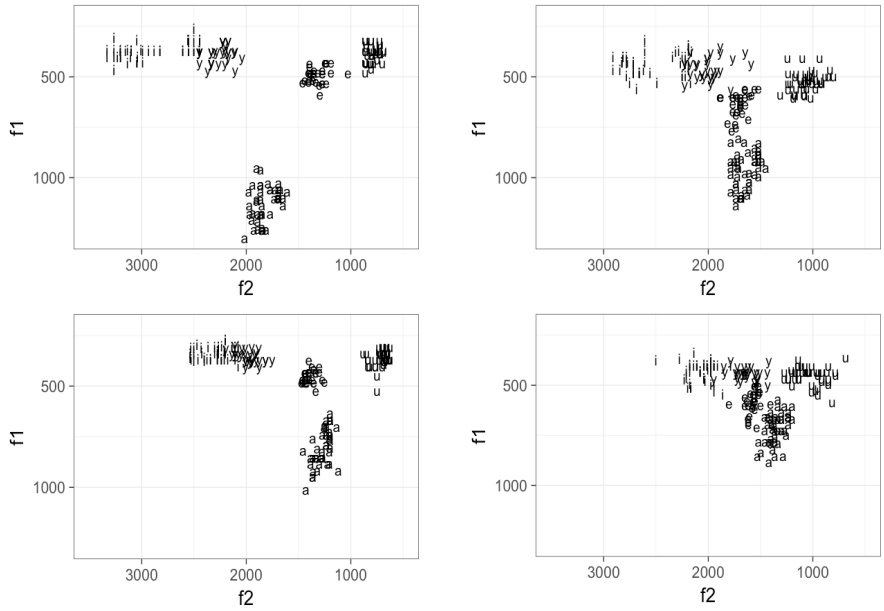


Figure 1. F1 and F2 of females (top) and males (bottom) with (right) and without (left) suffix.

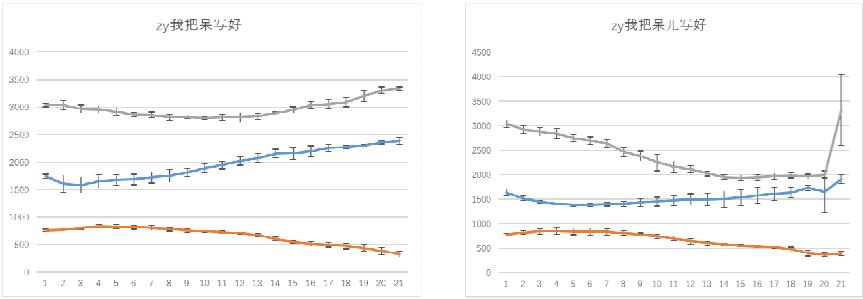


Figure 2. Formant trajectories for /ai/ without (left) and with the suffix (right).

Diphthongs and nasal codas

All four diphthongs /ai, əi, au, əu/ show a decrease in F3 from early on. When the suffix was added to the diphthongs ending on /i/ (/ai, əi/), F2 decreased whereas F1 did not seem to change (Figure 2). The suffix did not change F1 and F2 of the diphthongs ending on /u/ (/au, əu/, Figure 3).

On the other hand, the alveolar nasal coda /n/ deleted when the suffix was added, and the preceding vowel /a/ changed just like the monophthong /a/ (i.e., F1 decrease, F3 decrease). The velar nasal coda /ŋ/ was retained, but the rhotic suffix changed the preceding vowels, mainly by lowering their F3.

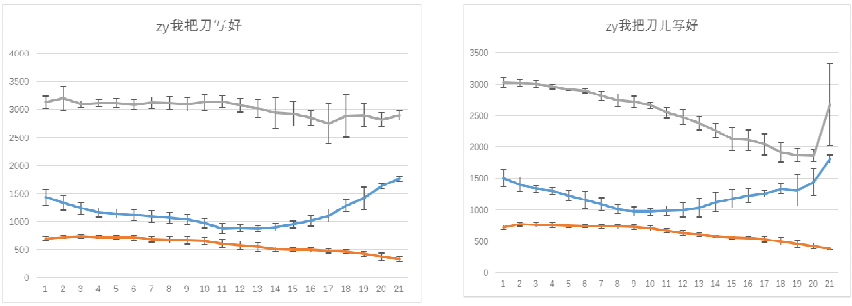


Figure 3. Formant trajectories for /au/ without (left) and with the suffix (right).

Conclusions

In summary, the rhotic suffix centralizes the preceding vowels in terms of F1 and F2 while lowers their F3. This study finds that the rhotic influenced all the monophthongs, unlike Huang (2010). The current results on diphthongs and nasal codas provide acoustic evidence for the previous phonological studies (e.g., Duanmu 2007). Future study will aim to investigate the underlying reasons for different behaviors of /i/-final and /u/-final diphthongs and alveolar and velar nasal codas.

References

Belotel-Grenié, A., Grenié, M. 1994. Phonation types analysis in standard Chinese, Proceedings of Spoken Language Processing pp. 343-346.

Boersma, P., Weenink, D. 2017. Praat: doing phonetics by computer.

Chao, Y.R. 1968. A Grammar of Spoken Chinese. University of California Press. Berkeley.

Duanmu, S. 2007. The Phonology of Standard Chinese (2nd Edition). Oxford University Press.

Huang, T. 2010. Er-suffixation in Chinese monophthongs: phonological analysis and phonetic data. Proceedings of the 18th International Conference on Chinese Linguistics (IACL-18). Vol 1, 331-344.

Lu, Y. 1995. Putonghua de Qingsheng he Erhua. Beijing, China: Shangwu Yinshuguan.

Tian, J. 2007 An Optimality Theory analysis of diminutive suffixation of Beijing Chinese, Proc. 23rd Northwest Linguistics Conference, Victoria BC CDA.

Zee, E., Lee, W-S. 2001. An acoustical analysis of the vowels in Beijing Mandarin. (pp.643-646). Aalborg, Denmark: Proceedings of INTERSPEECH.

Verb argument structure effects on tense: evidence from aphasia in Greek

Fani Nasika
School of Allied Health Professions, Canterbury Christ Church University, UK
School of Health, Aegean College, Greece

Abstract

This paper aims to present the findings of a study on tense in Greek-speaking individuals and whether it is affected by verb argument structure. Research evidence has shown that tense is impaired in aphasia but most of the studies concerned testing tense in isolation and not in dynamic contexts and in interaction with other factors. In this study, non-fluent and non-brain-damaged participants were tested on a two-choice sentence completion and a grammaticality judgment task. The results showed a significant difference in tense performance between the two groups but no overall significant effects of verb argument structure. However, there was individual variation within the non-fluent group which could be explained by the severity of aphasia, a processing or a tense deficit.

Key words: tense, verb argument structure, aphasia, Greek

Theoretical Background

The aim of the present study was to examine verb argument structure effects on tense in individuals with aphasia who are also native Greek speakers. Aphasia is an acquired language disorder caused by brain lesions in the language-dominant hemisphere which is usually the left hemisphere. A widely accepted aphasia classification system distinguishes two broad aphasia categories: non-fluent and fluent aphasia (Benson 1967). Non-fluent aphasia and, in particular, Broca's aphasia, is commonly related to a verb and tense deficit.

Tense is a grammatical feature and as such it denotes the time when an action has occurred. In languages such as English and Greek, tense is expressed through affixation, by combining a suffix with a verb stem. In these languages, Tense can be either [+PAST] or [-PAST]. Tense has been repeatedly found to be impaired in Broca's aphasia as in Kok et al. (2006), who observed that tense inflection was impaired in Dutch agrammatic individuals while similar diifuclties have been shown for Greek, too (Tsapkini, Jarema & Kehayia 2000; Varlokosta et al. 2006).

On the other hand, verb argument structure seems to affect verb production and processing in aphasia. Verb argument structure involves the subject of the verb and its complements. More precisely, a verb's argument structure may involve one argument, two arguments or three arguments,

which may be optional or obligatory. Moreover, the argument structure of some verbs involves argument movement. These verbs form the unaccusative category which is distinct from the unergative category.

Verb argument structure has been shown to be an influencing factor in aphasia. More precisely, verb argument structure complexity has been reported to affect verb processing in aphasia (Shapiro et al. 1987). Moreover, it has been claimed that two-argument verbs are processed faster than three-argument verbs (Shapiro, Gordon, Hack & Killakey 1993) while three-argument verbs seem more difficult to produce than two-argument verbs (Thomspon et al. 1997). Unaccusative verbs also seem more difficult to produce than unergative verbs (Thomspon 2003). The present study, thus, tested verb argument structure complexity effects on tense in Greek-speaking individuals with aphasia. The main hypotheses claimed that if verb argument structure complexity affects verb production and processing, then it might also be involved in the tense deficit often seen in aphasia.

Methodology

Sample

The study examined two groups of participants: 1) individuals with non-fluent aphasia and 2) non-brain-damaged individuals. The experimental group included four participants with non-fluent aphasia while the control group included six Greek native speakers. The participants in the control group had no neurological or sensory deficits. The participants in the non-fluent group were diagnosed based on the Greek standardized adaptation of the Boston Diagnostic Aphasia Examination (Papathanasiou et al. 2004).

Material and procedure

The testing procedure included two tasks. The first task was a two-choice sentence completion task and the second one was a grammaticality judgment task. The verbs included in these tasks were selected based on their high frequency, familiarity, and imageability. The first task consisted of 119 test items. The participants were presented with sentences which began with the time adverb *yesterday* in Greek but were missing their main verb. The participants had to complete the gap with one of two choices, with the verb either in the present or past tense. The grammaticality judgment task included 100 test sentences. The sentences were judged based on their verb argument structure or tense.

Results

The data collected from this procedure was analysed with multiple regression analysis. In the two-choice sentence completion task, the non-fluent group scored 65.13% correctly while the control group scored 99.72%. The two groups had a similar performance pattern regarding the number of arguments of the verbs. Performance was more accurate on verbs that had three optional arguments than two optional arguments, although this difference was not significant. In addition, the non-fluent group was more accurate when choosing the past tense for unergative verbs than for unaccusative verbs, but this difference was not significant either. Overall, in this task, the non-fluent group differed significantly from the control group in their ability to choose the past tense, irrespective of the argument structure of the verb. In the grammaticality judgment task, the non-fluent group's overall performance was lower than that of the control group (70.5% vs. 97.5%), with the first group showing a tendency to over-judge sentences as grammatical. This group, thus, showed a difficulty in judging ungrammatical sentences for their argument structure or tense. Correct performance varied across the two groups by the number of arguments of the sentences, with the non-fluent group judging two-argument structures more accurately than three-argument structures. Finally, the two groups differed significantly in their correct performance as far as grammatical and ungrammatical sentences were concerned.

Conclusions

Overall, this study has not found significant argument structure effects on tense in Greek aphasia, as in Koukoulioti and Stavrakaki (2014), while no significant difference was observed between unaccusative and unergative verbs. What was observed is the fact that the tense deficit in non-fluent aphasia is evident and does not concern only production. In addition, performance ranged within the non-fluent group regarding number of argument or argument structure complexity. The patterns seen comply more with the individualized profile of each participant and the level of the severity of their aphasia. In addition, processing parameters should also be taken into consideration since performance may have varied due to task demands or the processing abilities of each participant. Thus, the tense deficit seen in non-fluent aphasia should be interpreted in light of the severity of the aphasia, a processing deficit or tense specification difficulties so as to clarify clinical aspects for more effective clinical interventions.

References

Benson, D. F. 1967. Fluency in aphasia: correlation with radioactive scan localization. Cortex Vol. 3 No 4, 373-394.

Kok, P., Kolk, H., Haverkort, M. 2006. Agrammatic sentence production: Is verb second impaired in Dutch? Brain and Language 96, 243-254.

Koukoulioti, V., Stavrakaki, S. 2014. Producing and inflecting verbs with different argument structure: Evidence from Greek aphasic speakers. Aphasiology 28 (11), 1320-1349.

Papathanasiou, I., Feidatsi, M., Katsantoni, M., Panagiotopoulou, E., Malefaki, S. 2004. The validation of Boston Diagnostic Aphasia Examination (BDAE-3) in Greek. Paper presented at the 26th World Congress of the International Association of Logopaedics and Phonetics, 29 August-2 September 2004, Brisbane, Australia.

Shapiro, L. P., Zurif, E., Grimshaw, J. 1987. Sentence processing and the mental representation of verbs. Cognition 27, 219-246.

Shapiro, L. P., Gordon, B., Hack, N., Killackey, J. 1993. Verb-Argument Structure Processing in Complex Sentences in Broca's and Wernicke's Aphasia. Brain and Language 45, 423-447.

Thompson, C. K. 2003. Unaccusative verb production in agrammatic aphasia: the argument structure complexity hypothesis. Journal of Neurolinguistics 16, 151-167.

Thompson, C. K., Lange, K. L., Schneider, S. L., Shapiro, L. P. 1997. Agrammatic and non-brain-damaged subjects' verb and verb argument structure production. Aphasiology 11, 473-490.

aphasia: Evidence from agreement, tense and aspect. Aphasiology 20 (8), 723-743.

Using PRISM to profile semantic ability in Greek

Fani Nasika[1], Katerina Afantenou[2], Georgia Kremmyda[3]
[1]School of Allied Health Professions, Canterbury Christ Church University, UK
[2]School of Philosophy, Aristotle University of Thessaloniki, Greece
[3]School of Health, Aegean College, Greece

Abstract

This paper aims to show how PRISM-L and PRISM-G (Crystal, 1992) can be used to profile semantic ability or difficulties in Greek speakers. The two profiles were used for the analysis of three samples provided by one adult participant with aphasia and two normally developing children of 4;8 and 6;8 years old, respectively. The profiling procedure showed that PRISM can provided specific information regarding the range of the lexicon of an individual as well as details about the semantic analysis of their sentences. Based on PRISM, one can see the developmental course of a child and identify areas of clinical interest for further development or locate specific semantic difficulties at word and sentence level in adults with language impairments.

Key words: PRISM, semantic analysis, development, profile, Greek.

Theoretical Background

Profiles of linguistic ability or disability are a useful tool for clinicians since they provide a means for identifying areas of weakness and/or strength regarding language in populations with developmental or acquired language disorders. Profiles are very important since the clinician can use them for assessing the language abilities of an individual, make diagnostic hypotheses and plan focused and structured interventions based on specific linguistic deficits identified in these profiles. Crystal (1992) created profile charts which could describe linguistic abilities in distinct linguistic levels of analysis, such as phonology (PROPH), prosody (PROP), morphosyntax (Language Assessment, Remediation and Screening Procedure, LARSP) and semantics (PRISM). In 2016, several cross-linguistic adaptations of LARSP were published, including a Greek version, the Gr-LARSP, by Stavrakaki and Okalidou (2016). Within this context, this study aims to adapt PRISM (Crystal, 1992) in order to profile semantic abilities and difficulties in Greek-speaking individuals.

PRISM comprises two separate procedures. The first concerns the relationship between semantics and the mental lexicon (PRISM-L) and the second the relationship between semantics and grammar (PRISM-G), mainly referring to thematic-role structures. PRISM-L includes a chart with an inventory of 61 semantic fields, divided into 239 semantic sub-fields, which is used to identify the lexical range of an individual as well as the semantic

areas he/she uses. Interestingly, these semantic fields are organized in an acquisition order. In addition, PRISM-L identifies relational items (e.g., pronouns, prepositions, connectives, etc.) as well as paradigmatic and syntagmatic relations and developmental errors.

On the other hand, PRISM-G is structured on the basis of five developmental stages, on sequences of semantic elements (e.g., Actor, Activity, etc.) and on semantic relationships (e.g., addition, contrast, condition, cause, etc.). It offers the clinician the possibility to identify thematic-role structures across various developmental stages and also their range and complexity. Finally, PRISM-G describes the order-of-mention of clauses and whether this parallels the order of the external world.

In the present study, PRISM-L and PRISM-G were used to profile the linguistic abilities and/or difficulties of Greek-speaking individuals.

Methodology

Sample

The data was collected from three participants, one adult and two children. The adult participant was 37 years old, male, and was diagnosed with Broca's aphasia, after a stroke. The two typically developing children participants included one girl (4;8 years old) and a boy (6;8 years old).

Material and procedure

The procedure involved the collection of language samples from the three participants. These samples involved the narration of the Red Riding Hood story. The adult participant provided a written language sample of the story while the two children provided oral language samples. The samples were analyzed using the PRISM-L and PRISM-G procedures, following Crystal (1992).

Results

Regarding the written sample of the adult participant, PRISM-L showed that he used 86 content words and 52 functional words. The type/token ratio (TTR) for the lexical items was 0.61, which denotes a medium lexical range, while the TTR for the functional items was 0.41, which shows a small range for functional words. Regarding the content words, the adult participant used words from 24 out of 61 semantic fields, namely time, buildings, living, man, measurement, moving, making/doing, animals, clothing, sound, sight, happening, body, etc. Of these semantic fields, the ones with the smallest lexical range were man, world and the animals. It can be seen that these categories are spread over the acquisition continuum which is understandable since people with aphasia may lose access to parts of an

already acquired lexicon. Concerning PRISM-G, the adult participant produced mainly clauses of two or three semantic elements, which places him in Stages II and III. The most frequent combinations of basic functions and thematic roles that he produced were combinations of dynamic and stative verbs with temporal or locative semantic elements. Very often, he omitted the Actor of the action or the Experiencer or the thematic roles were reversed. Additionally, coordination and subordination were minimal.

On the other hand, the boy produced 53 function words, with a TTR of 0.3, and 66 content words, with a TTR of 0.6. In addition, he produced words from 19 out of 61 semantic fields while not producing words from the last 12 semantic fields of PRISM-L, which correspond to later stages of lexical development. The semantic fields covered were mostly man, moving, food, having, animals, thinking, feeling, clothing, furniture, make/do, size, body, building. Within these semantic fields, he produced words from various sub-fields, showing an evolving and developing vocabulary. Regarding PRISM-G, the boy produced clauses with four, three and two semantic elements, showing also coordination and subordination of two or more clauses. The main semantic relations presented at the clause level were addition, time, cause, location and purpose. The functions and thematic roles of the semantic elements were unambiguous and covered a wide range of combinations. According to Crystal (1992), the boy's profile matches Stage IV of PRISM-G and presents many elements of Stage V.

Finally, the girl produced 94 function words, with a TTR of 0.2, and 79 content words, with a TTR of 0.6, indicating a lot of repetition and a limited range of function words along with a wider range of vocabulary in the content words. The words produced by the girl included 18 out of 61 semantic fields, consisting mostly of the following: man, moving, quantity, procession, animals, food, language, sight. The semantic fields covered did not include the final 12 fields of PRISM-L, which are considered to be acquired later. As far as PRISM-G is concerned, the girl produces clauses with three semantic elements, with various combinations of thematic roles and functions. Moreover, she uses coordination extensively with 'and', linking together two, three or more clauses. Regarding subordination, it is infrequent while she uses only relative pronouns, such as 'which'. Thus, the profile of the girl matches Stage III, as far as clausal semantic elements are concerned, and presents a developing coordination and subordination system from Stage V.

Conclusions

The data analysis showed two different profiles: the one presented by the adult participant and the ones presented by the two normally developing children. Regarding the two children, it can be seen that the older the child,

the better they use function words and construct semantically complex sentences with subordination. As for the semantic fields acquired, it seems that the later fields of PRISM-L are not used extensively by young children and it may be the case that these lexical items are acquired and/or consolidated later. On the other hand, the adult profile showed elements of an already acquired lexical system, since he used words from various semantic fields, even from the ones found later in PRISM-L. However, the range of this vocabulary was medium which can be explained by the language impairment the adult has suffered due to his aphasia. In addition, his sentences contained mostly two or three semantic elements, which is lower than that of the 4;8 year-old girl, while he does not use complex sentence structures, either by coordination or subordination of clauses. This is also typical of individuals with aphasia and shows that PRISM-G can profile this difficulty.

To conclude, PRISM seems able to provide useful information on the development of semantic abilities in typically developing children and it can delineate the semantic profiles of individuals with aphasia, thus, providing an insight to their lexical range, semantic fields and their ability to form semantically sound sentences, which could inform clinical assessment and intervention.

References

Crystal, D. 1992. Profiling linguistic disability. London, Whurr Publishers.

Stavrakaki, S., Okalidou, A. 2016. Gr-LARSP: Towards a Greek version of LARSP. In Fletcher, P., Ball, M. J., Crystal, D. (eds.) 2016, Profiling Grammar: More Languages of LARSP. Multilingual Matters.

Syllable phonology and cross-syllable temporal production in Greek

Elina Nirgianaki[1,2], Athina Kontostavlaki[1], Olga Nikolaenkova[3], Maria Papanagiotou[2]

[1]Lab of Phonetics & Computational Linguistics, University of Athens, Greece
[1]Department of Primary Education, University of Athens, Greece
[3]Department of General Linguistics, Saint Petersburg State University, Russia

Abstract

The present study examined the durations of voiced vs. voiceless fricatives as well as their intrasyllabic and intersyllabic duration effects as a function of variable focus applications. In accordance with a production experiment, the results indicate the following: (1) voiceless fricatives are significantly longer than voiced fricatives; (2) the voice distinction of fricatives has both intrasyllable and intersyllable duration effects as well as compensation duration effects; (3) focus application is confined on a lengthening effect of the vowel in stressed syllable context.

Key words: consonant, vowel duration, syllable, temporal production, Greek

Introduction

The main aim of the present study is an experimental investigation of segment durations as a function of intrasyllabic and intersyllabic segment variability in variable focus contexts. In particular, the duration of voiced vs. voiceless fricatives and sibilants are investigated as well as their duration effects at syllabic as well as intersyllabic level.

Other context and prosody factors being equal, a great deal of studies with reference to segmental temporality support the following hypotheses: (1) segmental sounds and especially vowels have different intrinsic durations and, thus, “low” vowels, such as /a/ are intrinsically longer than “high” vowels, such as /i/ or /u/ (e.g. Fourakis, Botinis, Katsaiti 1999); (2) voiceless fricatives are longer than voiced fricatives at all places of articulation (Nirgianaki 2014); (3) both low/high vowels and voiced/voiceless fricatives have compensatory effects at syllable domain. (4) open syllables are correlated with longer nucleus vowels than closed syllables (Maddieson 1985). Furthermore, a variety of prosodic factors, including lexical stress, focus and speech tempo, have respective duration effects on different segmental sounds (e.g. Fourakis et al. 1999).

However, the cross-syllable temporal effects of segment variability have hardly been investigated and we thus attempt to document the temporal behaviour of segmental sounds at both intra-syllabic and inter-syllabic levels and outline basic similarities and dissimilarities between them.

Experimental methodology

The speech material consists of the four key words [ˈmasa] ‘chew’, [ˈmaza] ‘mass’, [ˈmiθi] ‘myths’, [ˈmiði] ‘mussel’ with lexical stress on the first syllable, produced at the context of the carrier phrase [ˈfonakse ___ ðinaˈta] ‘s/he said ___ loudly’. Five female speakers at their early twenties, with standard Athenian Greek pronunciation, produced the speech material at a normal tempo in a sound-treated studio at Athens University Phonetics laboratory. The speech material was analysed with Praat programme and segment, syllable, word and utterance duration measurements were taken. The results were subjected to statistical processing with SPSS statistical package.

Results

Three-way ANOVAs (voicing x place of articulation x focus) were carried out for utterances, normalized words, syllable and segment durations as the dependent variables.

First, in terms of utterance durations and normalized word durations (word/utterance), significant differences were noted only in terms of focus [$F(1, 112)=19.882$, $F(1, 112)=74.169$; $p<0.0001$], indicating longer words and utterances in focus condition. Normalized syllable and segment durations were then counted, representing the ratio of each syllable and segment duration, respectively, to word duration. The results are presented in table I and figures 1-4.

Table I. Mean normalized duration for each segment and syllable and mean absolute duration for each word.

Stimuli	C1	V1	S1	C2	V2	S2	Word (ms)
ˈmasa	0.250	0.317	0.567	0.264	0.169	0.433	353
ˈmaza	0.265	0.369	0.633	0.159	0.208	0.367	381
ˈmiΘi	0.277	0.299	0.575	0.186	0.173	0.425	363
ˈmiði	0.295	0.304	0.600	0.252	0.214	0.400	372

At the syllabic level, the results indicated [$F(1,112)=10.327$, $F(1, 112)=10.650$; $p<0.005$] that first syllable (S1) was significantly longer while the second one (S2) significantly shorter when the fricative was /z/ in non-focus condition.

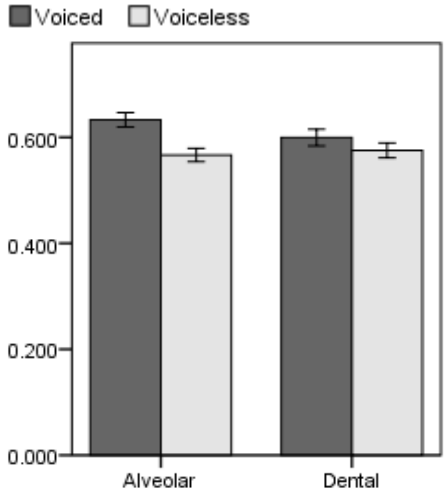


Figure 1. Mean normalized duration of first syllable as a function of fricative voicing and place of articulation.

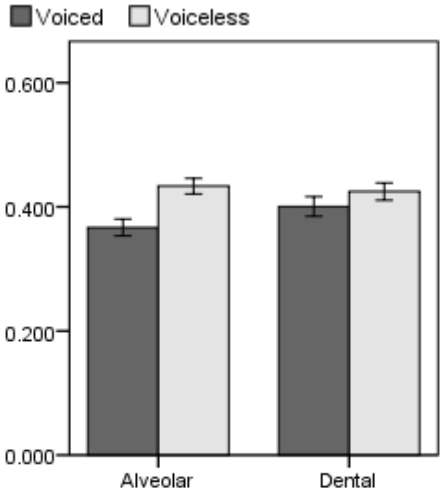


Figure 2. Mean normalized duration of second syllable as a function of fricative voicing and place of articulation.

At the segmental level, the first consonant (C1) was longer when the fricative was voiced and in non-focus condition [$F(1,112)=8.888$, $p< 0.005$], as well as when the fricative was dental [$F(1,112)=41.975$, $p <0.0001$]. The first vowel (V1) was significantly longer when fricative was /z/ [$F(1,112)=16.642$, $p<0.0001$], as well as in focus condition [$F(1,112)=12.650$, $p<0.005$]. Voiceless alveolar fricatives were significantly longer than voiced ones [$F(1,112)=12.551$, $p<0.005$] in non-focus condition [$F(1,112)=4.863$, $p<0.05$]. A main effect for voicing [$F(1,112)=51.435$, $p<0.0001$] revealed that second vowel (V2) was significantly longer when fricative was voiced.

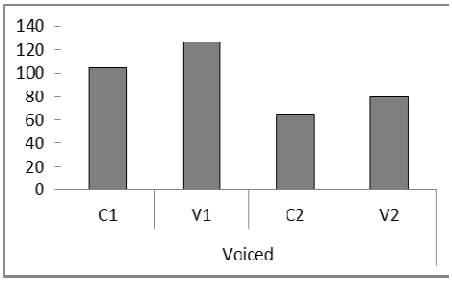


Figure 3. Mean normalized duration of segments as a function of fricative voicing.

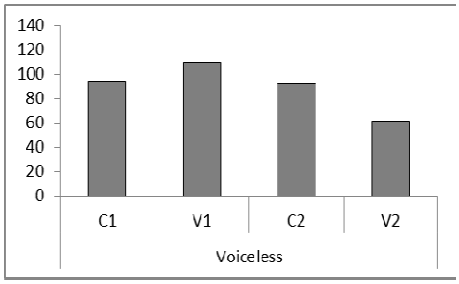


Figure 4. Mean normalized duration of segments as a function of fricative voicing.

Discussion and conclusions

The results of the present study indicate the following. In non-focus context, there seems to be a strong effect of fricative's voicing at both syllabic and segmental level, which is even stronger in the case of alveolar fricatives/sibilants. Regarding fricatives themselves, voiced fricatives – and mainly alveolars – are shorter than voiceless ones. When words contain voiced fricatives, all other segments are longer, while at the same time the duration of the word and utterance are not affected. Hence, the present results reveal that voiced/voiceless fricatives exhibit compensatory effects not only at intrasyllable level, but also at intersyllable and word levels, affecting previous segmental and syllabic durations as well. On the other hand, the temporal effect of focus is confined on the lengthening of the stressed syllable's vowel.

Syllable structure temporal correlates have mostly been associated with the nucleus vowel, according to which open syllables are correlated with longer vowels and closed syllables with shorter vowels (Maddieson 1985). This hypothesis has been corroborated in several studies but it still remain a controvertial issue (see McCrary 2004 for relevant discussion). The open syllable lengthening effect suggested by Maddieson (1985) has partly been corroborated in Greek along with a compensatory shortening effect on the intrasyllable onset consonant (Chaida, Dimoula, Magoula, Nikolaenkova 2017). The results of the present study, however, indicate that a variety of compensatory duration effects take place at both intrasyllable and intersyllable levels, which is a reflexion of different syllable structures along with a variety of other prosodic contexts.

Acknowledgements

Our thanks to Maria Papanogiotou and Athina Tsiori for data recordings. The Special Account for Research Grants, National and Kapodistrian University of Athens, has financed research and ExLing 2017 participation.

References

Chaida, A., Dimoula, I., Magoula, Nikolaenkova, O. 2017. Open vs. closed syllable phonology and temporal production in Greek (this volume).

Fourakis, M., Botinis, A., Katsaiti, M. 1999. Acoustic characteristics of Greek vowels. Phonetica 43, 172-188.

Maddieson, I. 1985. Phonetic cues to syllabification. In Fromkin, V. (ed.), Phonetic Linguistics, 203-221. London: Academic Press.

McCrary, K.M. 2004. Reassessing the role of the syllable in Italian phonology. Ph.D. thesis, UCLA.

Nirgianaki, E. 2014. Acoustic characteristics of Greek fricatives. Journal of the Acoustical Society of America, vol. 135, 2964-2976.

Modelling politicians' phono-styles with TTS

Carmen Patricia Pérez
CLILLAC-ARP, Université Paris Diderot – Paris 7, France

Abstract

An oratory technique often observed in politicians consists in playing with prosodic features in order to build a particular phono-style. In a previous perceptual and acoustic analysis, I showed that Hugo Chávez (former president of Venezuelan) and José Zapatero (former prime minister of Spanish) have completely different phono-styles, which are easily recognizable by listeners. In this study, text-to-speech (TTS) synthesis was initially used to produce neutral utterances. The prosodic parameters of these utterances were then varied to reproduce Chávez and Zapatero's specific phono-styles. The reproductions were tested in a series of perceptual experiments which show that the modifications made to intonation patterns are enough to identify the two politicians.

Key words: intonation, phono-style, public speech, politician speech, synthesis.

Introduction

A previous perceptive and acoustic analysis of Spanish-speaking politicians (Pérez 2016) revealed that H. Chávez (former president of Venezuela) and J. Zapatero (former prime minister of Spain) have completely different phono-styles in public 'spontaneous' speeches. In this study, I will first briefly describe Chávez and Zapatero phono-styles, then propose three perceptive tests using text-to-speech synthesis to convert written text into speech. F0, duration and chunking were modified to reproduce each politicians' prosodic patterns. The results of these tests show that Chávez and Zapatero are recognizable thanks to these stylized intonation pattern confirming the previous prosodic analyses.

Chávez and Zapatero prosodic pattern

Chávez and Zapatero's phono-styles were established thanks to the acoustic analysis of 22 utterances, extracted from public 'spontaneous' speeches. It was observed that each politician cuts his speech into chunks (1982ms for Chávez and 1127ms for Zapatero on average) and makes large pauses between chunks (811ms for Chávez and 720ms for Zapatero on average). Chávez's phono-style consists of a repetitive contour at the end of each intonation phrase (IP). This contour rises on the stressed syllable (the penultimate one for 80% of Spanish words), followed by a high dive of about 16 semi-tones on the last unstressed syllable of the word. Furthermore, we find lengthening on this final unstressed syllable (two, three or four times the length of the preceding stressed syllable). Zapatero's prosodic pattern is

characterized by small chunks and the use of a rising contour on the stressed syllable, continuing with a rise to a higher F0 value or floating tone on the last unstressed syllable. The politicians' schematized prosodic patterns are shown in fig.1 and 2.

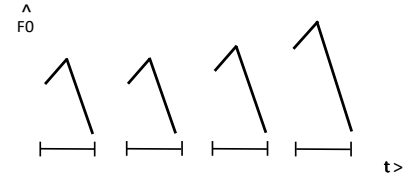


Figure 1. Chávez's prosodic pattern.

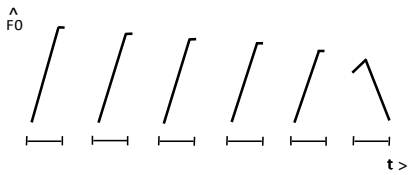


Figure 2. Zapatero's prosodic pattern.

Perceptual experiments: Methodology

Materials and design: The corpus consisted of four sentences adapted from the two politician's actual sentences, which were modified to be politically and dialectally neutral. They were produced using the TTS synthesis engine embedded in WinPitch. The utterances are 'prosodically neutral' (with a feminine voice) and were resynthesized according to the two politicians' respective prosodic characteristics. To ensure the prosodically modified speech was satisfactory, PSOLA generated pitch markers were carefully checked and edited.

In the first test, utterances were re-synthetized by modifying F0 parameter according to each of the above prosodic patterns. In the second test, F0 and duration of the last syllable of each chunk/IP were modified (the last unstressed syllable was twice as long as the stressed one). Each test had 12 synthetized utterances: 4 Chávez (C), 4 Zapatero (Z) and 4 'normal' (N) produced by TTS. Finally, in the third test, Chávez's utterances from the second test and Zapatero's utterances from the first were used. Four utterances were also added with shorter chunks for each politician. This test thus had 20 utterances: 4 C, 4 Z, 4 N, 4 C with shorter chunks (CSC) and 4 Z with shorter chunks (ZSC). Speed and pauses of the tests were not modified.

Participants: 10 native speakers of Spanish (Spanish or Venezuelan) and 10 non-Spanish-speaking participants (Chinese or Japanese mother tongue) who would, in a certain way, give a psychophysical (and not linguistic) perception.

Procedure: Participants started with a learning task of five minutes where they first listened to the unmodified Chávez and Zapatero recordings in order to get accustomed to their phono-style, followed by synthesized utterances in order to get used to synthetic female voice. Finally, they listened to re-synthetized Chávez and Zapatero utterances in order to test the phono-styles' recognition. Each test was based on a seven-step answer scale: surely C, probably C, maybe C, Other, maybe Z, probably Z and surely Z. Each step corresponded to a number: -3, -2, -1, 0, 1, 2, 3 respectively.

Results

Test 1 shows that Chávez is not well recognized; Normal (Other) and Zapatero may be recognized, but most of the time confounded: the recognition variability is considerable (cf. SD on fig.1). I conclude that the F0 pattern alone is not enough to enable the speakers' differentiation.

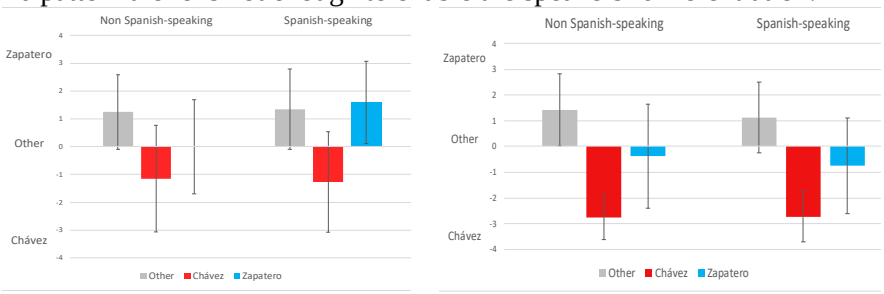


Figure 3. Test 1 (left) and 2 (right) results. Mean and SD of politicians' recognition.

Test 2, where the utterances had a lengthening of the IP last syllable, shows that Chávez is very well recognized; he is never confused with Zapatero or with normal utterances (cf. fig. 4). The normal condition is confused with Zapatero, while Zapatero is often confused with Chávez. I conclude that lengthening is a characteristic of Chávez and not of Zapatero.

Test 3 shows that the Chávez pattern is still very well recognized despite the shorter chunking. On the contrary, the shorter chunks help the Spanish native speakers to recognize Zapatero. Normal utterances are confused with Zapatero's but the mean is the smallest one of the three tests.

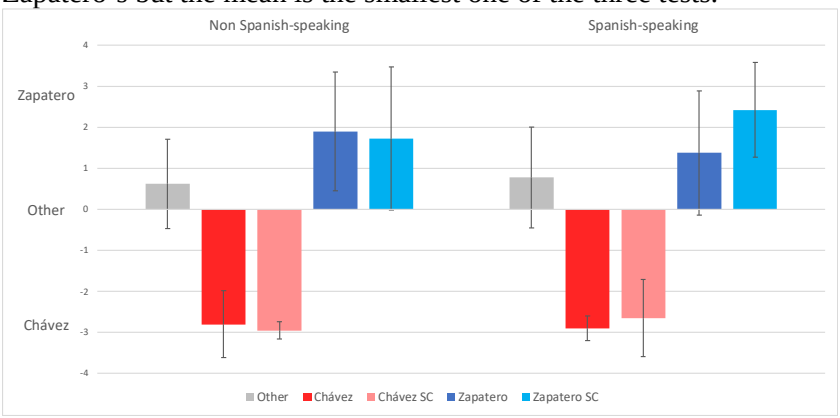


Figure 5. Test 3 results. Mean and SD of politicians' recognition.

Conclusion

These synthesis experiments show that in general F0 modification is not sufficient to identify Chávez, who is well recognized only when the last IP syllable duration has been raised (test 2 and 3). Zapatero's pattern is often confused with the 'normal' one, probably because his prosodic pattern is similar to the 'standard' phono-syntactic Spanish intonation. Shortening the chunks provides very good Zapatero recognition by Spanish-speaking hearers. Results illustrates that I) the F0 movement and the duration of the last unstressed syllable is crucial in these politicians' identification and for oral Spanish and II) Chávez and Zapatero phono-style are easily recognizable regardless participants mother tongue.

References

Fónagy, I., 1983. La vive voix. Essais de psycho-phonétique. Paris: Payot.

Léon, P., 1993. Précis de phonostylistique. Parole et expressivité: Nathan.

Martin, P., 2010. Intonation in Political Speech: Ségolène Royal vs. Nicolas Sarkosy. Rome, pp. 54-64.

Martin, P., 2015. The structure of spoken language. Intonation in Romance: Cambridge University Press.

Pérez, C.P., 2016. A study of the phono-styles used by two different Spanish-speaking political leaders. Proc. of Speech Prosody 8, Boston, pp. 410-414.

Sosa, J. M., 1999. La entonación del Español: Su estructura fónica, variabilidad y dialectología. Madrid: Catedra.

Touati, P., 1995. Pitch range and register in French political speech. Proc. XIII International Congress of Phonetic Sciences, Volume 4, pp. 244-248.

Prosodic contours of migrant learners of French

Darya Sandryhaila-Groth
UFR Linguistique, LLF, Paris-Diderot Paris 7, France

Abstract

This study concerns an aspect of L2 learning: the realization of the prosodic contours of short French sentences read by migrant learners. They are enrolled in an intensive French course at advanced A1 level as defined by the CEFR. The multimedia software WinPitch is used for analyses.

Key words: L2 in immersion, prosodic contours, prosodic morphing.

Introduction

In France, according to the national law, immigrants need to learn French in a language school as a part of integration into French culture and daily life. To stay in France and to pursue the professional studies, they have to pass a proficiency exam called "Diplôme Initial de Langue Française" (DILF - *Initial French Knowledge Diploma*) (*Gloaguen-Vernet 2009*). This exam is only awarded to foreigners living on French territory. To help the learners improve their oral expression, we worked on the correction of their pronunciation of French sentences, needed to pass DILF oral exam.

Methodology

For this study, two recordings of learners of French were analyzed: four male and three female young adults, without any prior knowledge of French and with different L1 (Bangla, Portuguese, Albanian and Chinese). They completed three phases: a first and a final recording and a training phase with WinPitch LTL (Language Teaching and Learning) in between.

Hypothesis

Hypothesis 1: during the first recordings, the learners will do much more chunking while reading, as a lack of syntactic and prosodic knowledge of sentence grouping and a weak knowledge of the French graphic code. Hypothesis 2: during the final recordings, the learners will keep the correct prosodic pattern, because the real-time visualization during the training phase with WinPitch LTL will allow them to achieve a good cohesion of text, including the prosodic pattern, thanks to the matching of the prosodic structure and the syntactic one.

Corpus

Six declarative sentences were recorded and analyzed. The final realization was compared with the model (native French speaker). The model speaker had to read the sentences rather slowly as requested for A1 learners, but emotional variations were still realized.

Intonational model

According to the intonational model of P. Martin (Martin 1975, 1982, 2015), every prosodic contour has a different movement on stressed syllables depending on the contour of the last word in the sentence. There is a hierarchical order of the syntactic units. The prosodic structure is based on the slope contrast between successive stressed (prominent) syllables. At the first level, rising contours are called C1 and falling ones C0. In the inventory of contours, there can be furthermore a C2 and a Cn. Within a single sentence there is the same amount of intonational contours and accented syntactic units (Lepetit and Martin 1990). According to the phonosyntax principle, the prosodic structure should match the syntactic one. All recorded sentences were annotated following this principle.

An example of a sentence analyzed with WinPitch

In Figure 1, left panel, the model speaker read: Mart**i**ne met ses l**i**vres et ses cah**ie**rs sur la t**a**ble. "*Martine puts her books and her exercise books on the table.*" There are four prosodic contours, located on stressed vowels, marked in bold in text: C2 green, Cn blue, C1 maroon, and C0 yellow. In right panel, a male learner realization of the same sentence in the final recording. He realized the same contours and a similar duration as the model.

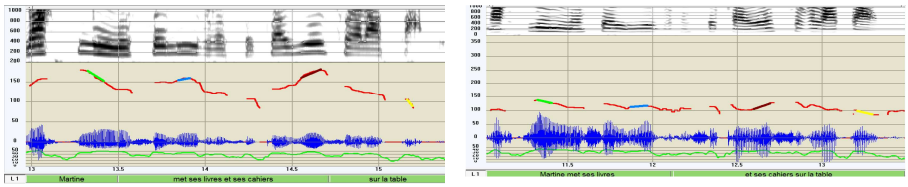


Figure 2. Left: Contours of the model. Right: final recording of a learner.

Results

Figure 2 summarizes the learner's production of all corpus sentences and their contours. The blue bars represent the realized number of contours for each of the seven learners during the first recording. In this first phase, the learners didn't hear the model yet and read the sentences according to their own understanding. The red bars represent the realized number of contours

during the final recording. In this last phase, we compared the learners' contours with the model.

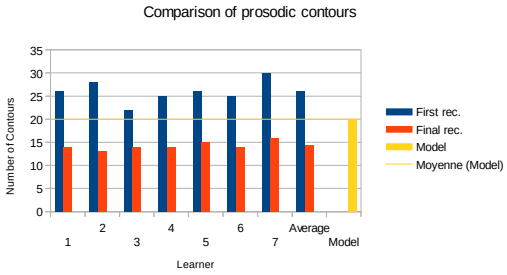


Figure 3. Comparison of prosodic contours between learners' first and final recordings, and the model speaker.

Next in Figure 3, on the y-axis, more detail of the learners contours vs., the model contours are on the x-axis. The learners realized all final falling contour C0 correctly. The rising contour C1, the falling contour C2, and the neutral contour Cn which can slightly rise or fall or stay at the same melodic level. Cn was the most problematic contour. I also add a C+ to show that the learners realized a contour, while no one exists in the model. Moreover, a Ø-sign indicates when the model didn't stress a grammatical word, while the learners stressed the word.

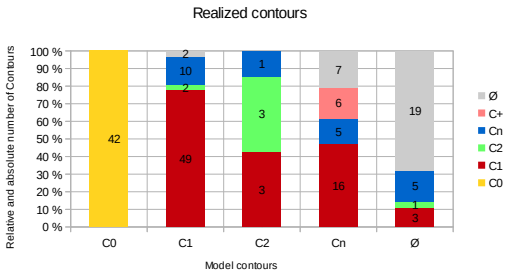


Figure 4. Number of realized contours in final recording. Relative numbers on y-axis, and absolute numbers as inset.

Finally in Figure 4, the comparison of sentences duration is shown. During the first recording, the learners realized a lot of chunking and stressed nearly each word (Guimbretière 1997). The duration of the final recording almost matches with the model duration, while the duration of the first recording is 1.4 times longer on average.

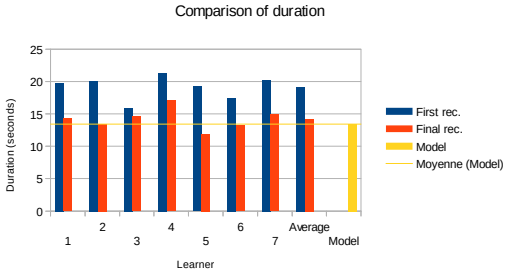


Figure 5.Comparison of sentence duration between first and final recordings.

Prosodic morphing

Figure 5 shows an example of prosodic morphing for the initial example given in Figure 1. During the first recording of the sentence, left panel, the learner realized a C1 rising contour instead of a C2 falling one. The prosodic morphing was done with the help of WinPitch: F0 modifications are shown in dark green and blue, pauses in yellow, intensity in light blue and duration in purple (right panel). With this manipulation, the learner can hear his own voice producing the right movement. He improved his realization of the contours and the sentence duration was much shorter.

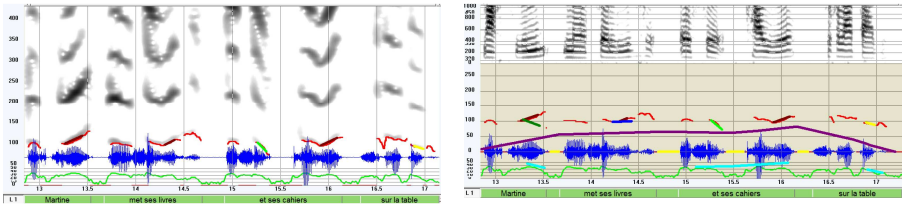


Figure 6. Left: first production of learner. Right: prosodic morphing of the same production.

Conclusions

The analyses of the first recording show that the first hypothesis is confirmed: the learners produce more chunking while reading unknown sentences. The final recordings support the second hypothesis, that the learners were able to shorten the duration of the sentence and they achieved a better cohesion of text, although some mistakes in the contours remain.

References

Gloaguen-Vernet, N. 2009. Enseigner le français aux migrants. Paris, Hachette Livre.

Gumbretière, E. 1997. Analyse des faits prosodiques dans une perspective d'enseignement/apprentissage. Polyphonie pour I. Fonagy, par J. Perrot, l'Harmattan.

Lepetit, D., Martin, Ph. 1990. Étude différentielle intonative français/anglais. IRAL, vol. XXVIII/2.

Martin, Ph. 2015. The structure of spoken language: intonation in Romance. Cambridge University Press.

Martin, Ph. 1982. Phonetic realizations of prosodic contours in French. Speech Communication I, 283-294.

Martin, Ph. 1975. Analyse phonologique de la phrase française. Linguistics 146, 35-67.

www.winpitch.com

The dynamics of prosodic adaptation between Italian conversational partners

Michelina Savino
Dept. of Education, Psychology, Communication, University of Bari, Italy

Abstract

This paper explores the dynamics of prosodic adaptation between speakers of Bari (BI) and Lecce (BI) varieties of Italian, in which polar question intonation contour is different as to the assignment of the rise for marking questioning: on the nuclear syllable in BI, on the phrase boundary in LI. Intonation analysis and distribution of contour types produced by BI-LI speaker pairs in a game-based interaction revealed that only BI speakers tended to accommodate their question intonation to their LI partners'. Dynamics of adaptation varied across BI speakers, hinting to different accommodation strategies which not necessarily entail overriding the rule of their native variety intonational grammar in spontaneous imitation.

Key words: Prosodic adaptation, question intonation, varieties of Italian.

Introduction

Research on prosodic accommodation in verbal interaction has been mainly concentrated on quantitative measurements of acoustic-prosodic features (e.g. Levitan et al. 2015), whereas little attention has been paid to intonation contours. This paper focusses on the dynamics of prosodic adaptation by analysing the F0 contour of polar questions produced by Bari Italian (BI) and Lecce Italian (LI) speakers. This case is particularly interesting since the prototypical yes-no question tune is completely different in the two varieties as to the assignment of the functional rise for marking questioning (Savino 2012): BI has the rise on the nuclear syllable followed by either a falling or a rising terminal (L+H*L-L% and L+H*L-H%, respectively), whereas LI has the rise on the phrase boundary preceded by a low or falling nuclear accent (L*L-H% or H+L*L-H%). Therefore, in case of interacting BI-LI speakers pairs spontaneous imitation of the conversational partner's question contour would entail overriding the rule of the native variety intonational grammar as to the assignment of the rise for marking interrogativity. For BI speakers, an alternative strategy which would not imply overriding such a rule could consist in the selective imitation of the L-H% boundary, thus preserving the realisation of the nuclear rise (resulting in a prevailing production of L+H*L-H% over L+H*L-L%). Having this as a background, this study explores the dynamics of prosodic adaptation in pairs of BI and LI speakers asking polar questions during a game-based interaction.

Method and materials

Yes-no questions were elicited by involving pairs of BI-LI participants in the "Guess who?" game, where each participant is provided with a board having a set of pictures of fe/male characters drawn on it. The game starts with each player selecting a card from a separate pile of cards with the same characters drawn on the board, and it consists in being the first to guess which card the partner has selected, by asking exclusively yes-no questions on characters' features in order to eliminate candidates. In each recording session, participants were sitting in front of each other, wearing an AKG C520 condenser microphone headset connected to a Marantz PMD 661 digital recorder. Eye contact was inhibited, and each recording session lasted 35 min on average.

Informants

A total number of 10 speakers (5 from Bari and 5 from Lecce) participated in the recording sessions. They were all female students at the University of Bari, aged 22-24 and not familiar with each other. Gender, age and familiarity were controlled as parameters since they are been found as relevant in accommodation processes (e.g. Pardo 2006).

Data annotation and measurements

Around 1,000 polar questions (around 100 for each speaker) were collected and manually annotated by using Praat (Boersma 2001), along these levels:

- segments (utterances, nuclear words and syllables)
- tunes (nuclear pitch accents and boundary tones, AM framework).

The following measurements were performed:

- relative frequencies of contour types produced by BI and LI speakers in each pair/session (direction of adaptation);
- relative distribution of contour types in the first vs second half of the dialogue (dynamics of convergence at the overall dialogue level)
- distribution of imitated vs not imitated F0 contour in BI-LI adjacent turns where question is also textually repeated (adaptation dynamics at the turn-level, influence of lexical imitation on tune imitation). Due to space limitations, results for this measure will not be presented here.

Results

The relative frequencies of all contours types produced by BI and LI speakers in each pair were calculated, as illustrated in Fig.1. It can be noted that, in all sessions, LI speakers show a clear tendency to not divert from using their prototypical (falling-)rising contour when asking questions, whereas BI speakers exhibit a larger variability in their question tune

realisations. This indicates that prosodic adaptation was asymmetric, i.e. only BI speakers tended to accommodate their intonation contour to their LI conversational partners'. Only one LI speaker (LI_sp1 in Fig.1) shows an amount of rising-falling contours in her productions, but this contour type is different in tonal alignment with respect to the prototypical BI L+H*L-L% . In fact, it is phonologically described as L*+HL-L% (i.e. with a nuclear peak later that in BI), and reported by Savino (2012) as a non-prototypical question tune in LI.

Accommodation behaviour vary across BI speakers: one speaker (BI_sp1 in Fig.1) does not show to adapt intonationally to her LI partner, since her questions are characterised exclusively by the BI nuclear L+H* followed by either an L-L% or L-H% boundary, with L-H% prevailing over L-L%. The remaining four BI speakers all exhibit a certain amount of the typical LI (H+)L*L-H% contour: BI_sp2 and BI_sp4 both have 60% of their polar questions realised with a (fall-)rise instead of their native rise-fall or rise-fall-rise, whereas BI_sp5 and BI_sp3 exhibit (fall-)rises in less than 30% of the questions asked to their LI partners. As to their native BI question contour, also for these four speakers the production of L+H*L-H% prevails over L+H*L-L%. This outcome, common to all BI speakers, could be interpreted as a selective imitation of the LI (H+)L*L-H% contour.

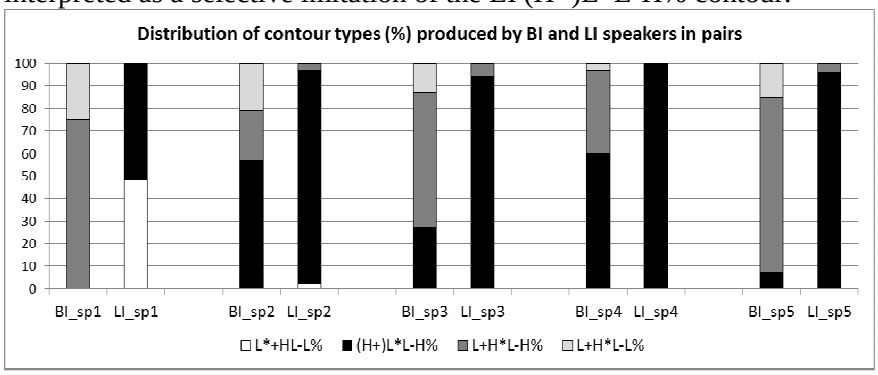


Figure 1. Distribution of question contour types (in perc.) produced by Bari Italian (BI) and Lecce Italian (LI) speakers.

Since prosodic accommodation was asymmetric, the relative frequencies of the question contour types as distributed in the first vs second half of the dialogue were determined only for BI speakers, in particular for speakers BI_sp2-3-4. Results in Figure 2 reveal two different types of accommodation dynamics across these speakers. For BI_sp3, the production of (H+)L*L-H% contours is almost all concentrated on the second half of the game session, this implying a gradual prosodic convergence (via a underlying gradual imitation process) towards LI partner's question intonation. Instead, BI_sp3

and BI_sp4 started realising the typical LI question contour early in their interaction, showing a rather uniform distribution of (H+)L*L-H% between the first and the second halves of the dialogue. The latter behaviour seems to suggest for BI_sp3 and BI_sp4 a larger variability in their question intonation repertoire which could include the (fall-)rise along with the prototypical rise-fall and rise-fall-rise F0 contours: when interacting with a LI conversational partner, the (fall-)rise contour can be easily "retrieved" early in the dialogue for prosodic adaptation purposes.

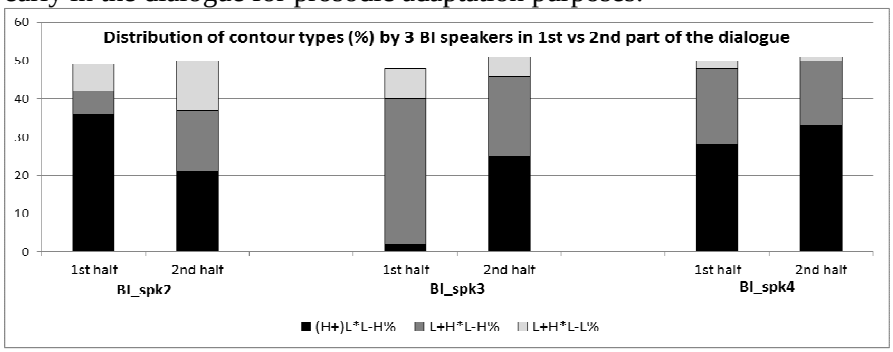


Figure 2. Distribution of question contour types (in %) produced by BI speakers 2-4, as occurring in the first vs the second half of the game session.

Summary and conclusions

We analysed the F0 contour of polar questions produced by BI and LI speakers pairs involved in a game-based interaction and found asymmetric adaptation, in that only BI speakers tended to accommodate their question intonation contour to their LI conversational partners'. Dynamics of prosodic adaptation can vary across BI speakers, hinting to different accommodation strategies which not necessarily entail overriding the rules of the native variety intonational grammar in spontaneous imitation of the conversational partner's question contour.

References

Boersma, P. 2001. Praat, a system for doing phonetics by computer, Glot International 5(9/10), 341-345.

Levitan, R., Benus, S., Gravano, A., Hirschberg, J. 2015. Acoustic-prosodic entrainment in Slovak, Spanish, English and Chinese: A cross-linguistic comparison. In Proc. of the SIGDIAL 2015 Conference, 325-334, Prague, Czech Rep.

Pardo, J. 2006. On the phonetic convergence during conversational interaction. JASA 119(4), 2382-2993.

Savino, M. 2012. The intonation of polar questions in Italian: Where is the rise? Journal of the International Phonetic Association 42(1), 23-48.

Long vowels in Mongolian

Yumei Sang
ARP, UFR Linguistique, Université Paris Diderot Sorbonne Paris Cité, France

Abstract

This paper investigates one of the fundamental and controversial questions in Halh Mongolian: the phonological status of long vowels (VV), especially VV in non-initial syllables and the distribution of VV. This study questions the traditional analysis of vowel quantity based on the single criterion of the realisation of vowel length, which is purely phonetic. This paper takes into consideration the historical development of VV in Mongolian. It demonstrates that, synchronously, the distinction between long (VV) and short vowels (V) should consider the distinct behaviours that reveal the different nature of these two types of vowels.

Key words: Mongolian vowel, long vowel, vowel distribution, vowel devoicing

Introduction

Mongolian is a language with contrastive vowel quantity. The seven long vowels [aː], [ɔː], [ʊː], [eː], [oː], [uː], [iː] contrast with the seven short vowel [a], [ɔ], [ʊ], [e], [o], [u], [i] by a difference in length. However, an important variation of vowel duration is observed both for short and long vowels. This variation of vowel length is the source of different theories with regards to VV and their distribution. According to a more traditional point of view (Damdinsuren 1959, among others) VV have free distribution and all VV are phonologically long regardless to their position within a word. According to Svantesson *et al.* (2005), long vowels only appear in the initial position of a word and non-initial VV are treated as short vowels. This later point of view is based on vowel duration.

This present study examines several aspects of acoustic properties of long and short vowels in order to determine the phonological status of non-initial VV sequences and the distribution of long vowels. In addition to the analysis of vowel length and vowel formant structures, a careful study of vowel devoicing is conducted on two types of corpus: target words (controlled speech) and texts (continuous speech). The analysis of vowel devoicing is crucial in understanding the phonemic status of vowel quantity.

Before starting the analysis, the historical development of long vowels in Mongolian should be given briefly. There are two opposite points of view regarding the origin of long vowels. On the one hand long vowels are considered as the result of the merger of two short vowels (Sanzheyev 1953). Roughly, in VCV sequences, the intervocalic consonants (ɣ, g, ɣ, ɓ, ɦ) dropped and the two short vowels merged into one long vowel. On the other

hand, it is believed that long vowels in Mongolian are primary (Poppe 1962). However, according to these two theories, long vowels appear in all positions within a word and VV initial and non-initial are developed from the same process.

Methodology

The data consist of 600 target-words embedded in a frame sentence [pi__ dedʒhelʒsen] 'I__ that said, (I said that__) read by six native Mongolian speakers. In total, 1655 long and short vowels are analysed using the signal processing software WinPitch. The segmentation of vowels and consonants is performed manually and visually. The segmentation method is based on oral constriction (Turk *et al.* 2006). The beginning and the end of F1 and F2 are taken as vowel duration. For vowel devoicing, additional 4 texts containing 1200 words and 2079 long and short vowels are analysed. For the analysis, VV sequences are divided into initial VV and non-initial VV.

Analysis and results

Measure of vowel length

The duration of vowels is measured. The average length, standard deviation and median values are given in Table 1. :

Table 1. Average length, standard deviation and median values for initial VV, non-initial VV and V.

	Initial VV	*Non-initial VV*	*V*
Average	183	127	75
Stan. deviation	35	33	22
Median	182	124	75

This table shows that the average lengths are very similar to median values. This indicates the small dispersion of the data. The average means for initial VV, non-initial VV and V are 181 ms, 127 ms and 75 ms respectively. In percentage, short vowels thus represent 42% of the duration of long vowels and non-initial VV are 69% of the duration of long vowels. Non-initial VV are exactly in the middle between long and short vowels in terms of duration as shown in Figure 1 below:

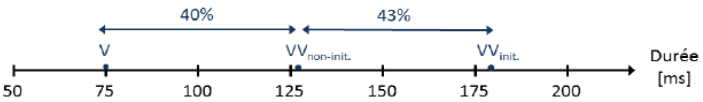


Figure 1. Distance between non-initial VV and V, distance between non-initial VV and V.

There are thus no reasons to consider that non-initial VV are short rather than long, and vice versa. It also reveals that vowel duration alone cannot constitute a solid factor to determine the phonemic status of vowels, especially for non-initial VV sequences.

Measure of vowel formant structure

There is evidence from several languages that vowel length impacts vowel quality (Lindblom 1963). Long vowels display more peripheral vowel quality whereas vowels with shorter phonetical length show more centralized vowel triangle. Figure 2 shows the result of formant analysis in Mongolian:

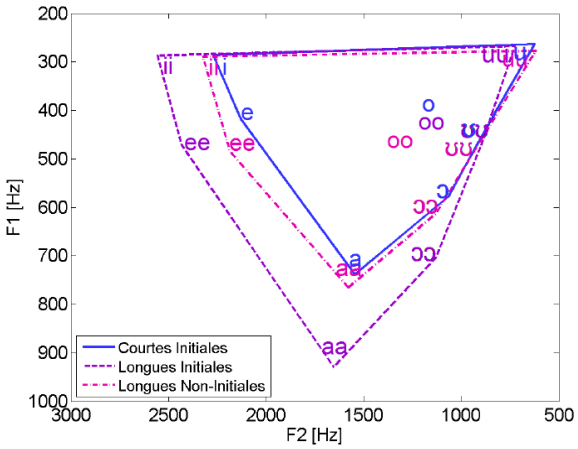


Figure 2. Vowel triangle for initial VV, non-initial VV and V.

Figure 2 shows that in Mongolian, vowel length does have an impact on vowel quality. Nevertheless, the correlation between vowel quantity and vowel quality seems to depend on each vowel individually rather than on vowel type. The series of [u], [ʊ] and [o] show little difference in vowel quality and some even overlap. This result shows that vowel quality also does not provide a clear indication for the phonological status of non-initial VV.

Vowel devoicing

The analysis of vowel devoicing here follows the cross linguistic survey of vowel devoicing by Chitoran and Marsico (2010). Vowel devoicing is a process in which vowels are produced with open glottis. Vowel devoicing can be caused by aerodynamic factors (Ohala1983) or/and by the glottal gestural overlap between voiceless consonants and short vowels (Jun and Beckman 1993). The results of the analysis show that V is highly prone to devoicing, both partially and completely. No case of vowel devoicing is

observed either for initial and non-initial VV (Sang2016). Figure 3 show one example of vowel devoicing of [i] in [xiʧeeʖ] (lesson):

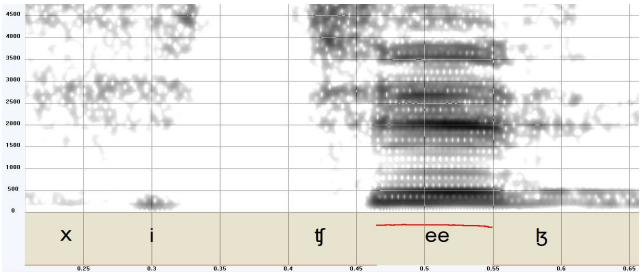


Figure 3. Complete vowel devoicing of [i] and [ee] without devoicing

Initial and non-initial VV behave identically in terms of vowel devoicing and V differs from them. The identical behaviour of initial and non-initial VV is governed by their identical underlying status. The longer a vowel, the less it is affected by devoicing.

Conclusion

This paper provided evidence that vowel length alone cannot be a reliable factor for a distinction between long and short vowels. Other behaviour patterns, such as vowel devoicing must be taken into consideration. It is concluded that initial VV and non-initial VV are both long and therefore long vowels have free distribution in Mongolian.

References

Chitoran, I., Marsiko, E. Vowel devoicing – an updated phonetic typology. 36th Meeting of Berkeley Society, Berkeley, CA, 2010.

Damdinsuren, St. Orchin tsagiin mongol hel zui. Ulaanbaatar, 1996.

Jun, S.A., and Beckman, M.E., A gestural-overalp analysis of vowel devoicing in Japanese and Korean. 67th LSA Annual Meeting. Los Angeles. CA, 1993.

Lindblom, B. Spectrographic study of vowel reduction. JASA, 35(11):1744-1781, 1963.

Ohala, J.J., The origin of sound patterns in vocal tract constraint, in: The production of speech. New York: Springer-Verlag.189-216, P.F.MacNeilage(ed), 1983.

Poppe, N. The primary long vowel in Mongolian. JSFOu, 63(2):1-19, 1962.

Sang, Y. Topics in Mongolian Phonetics and Phonology. Ph.D. thesis. Paris, 2016.

Sanzheyev, G.D., Spavnitelinaya grammatika. Moskva, 1953.

Svantesson, J.-O., Tsendina, A., Karlsson, A., Franzen, V., The phonology of Mongolian. Oxford UniversityPress, 2005.

Turk, A., Nakai, S., Saguhara, M., Acoustic segment duration in prosodic research: A pratical guide. 2006.

Methodological challenges in investigating nominalized infinitives in Spanish

Barbara Schirakowski
Freie Universität Berlin, Germany

Abstract

The present paper deals with Spanish nominalized infinitives (NI), a type of syntactic nominalization that comes in different subtypes. On the basis of argument realization, I will focus on two different forms, a more nominal and a more verbal one. Previous research has mainly focused on the internal syntax of NI, while the factors that govern the distribution of different subtypes have remained unclear. In the following, I will discuss some methodological challenges in investigating the latter issue and suggest acceptability judgment tasks as a possible solution. The results of two experiments show how the acceptability of NI is determined by the interaction of two factors: 1) argument realization and 2) event interpretation.

Key words: acceptability judgments, nominalized infinitives, Spanish

Introduction

Spanish allows a number of different NI, i.e. infinitives that are preceded by a determiner. For the present purpose, I will focus on only two forms and distinguish them on the basis of argument realization. Type A is the more nominal form and maps an argument of the verbal base onto a PP introduced by the preposition *de*, cf. (1). Type B is the more verbal form that selects a direct object, which is often, but not necessarily, a bare noun, cf. (2).

(1) el cantar de los pájaros
'DET singing of the birds'

(2) el cantar (las) coplas
'DET singing (the) songs'

Questions regarding the distribution of different subtypes and possible interpretations of NI have received little attention in the literature so far. The aim of this study is to bridge some gaps in this domain by presenting results from acceptability judgment experiments. At first, I will provide some necessary background information and then proceed to the methodology and the experimental results.

Background

Type A has received more attention in the literature, especially with regard to possible verb classes (e.g. Demonte & Varela 1997). It has been noted that it can be built quite freely on intransitive verbs and that it is dispreferred with transitive ones. It is not entirely clear whether an argument of a

transitive verbal base can be mapped onto the PP position. Realization of the Theme argument is often considered 'ungrammatical' and realization of the Agent is rarely discussed.

From a semantic point of view, it is noteworthy that both type A and type B can occur as event-denoting nominals. More recently, the literature has become interested in the distinction between episodic and generic readings of deverbal nominalizations, cf. (3) vs. (4), and it has been noted that some (more) verbal forms show a preference for generic event interpretations (Iordăchioaia & Soare 2015 among others).

(3) The observation of the neighbors *was completed at midnight*.
(4) The observation of the neighbors *is a bad habit*.

The central question of this paper can be formulated as follows: How does the interaction of argument realization (Theme vs. Agent) and event interpretation (episodic vs generic) influence the distribution of type A vs. B?

Methodology

Methodological considerations

Difficulties in answering this question result from different factors. Spanish has a variety of derivational affixes that are highly productive and NI are by no means the default strategy for deverbal nominalization. They are often criticized in terms of style and usually occur in certain registers typical of written language. This background makes it difficult to reach reliable generalizations about their acceptability. Furthermore, to evaluate the impact of different interpretations, its seems essential to view NI in unambiguous contexts that evoke episodic or generic interpretations. Contexts that are available in corpora usually do not fulfil this requirement, but show all kinds of ambiguities. An experimental approach seems suitable to overcome these problems. Firstly, acceptability judgment experiments allow the NI to be embedded under matrix predicates that allow for only one interpretation. Secondly, judgments can be elicited on a multi-point scale which makes it easier to detect subtle differences between structures that are rather marginal, but not "equally bad" (e.g. Featherston 2007).

Experimental design

The two experiments that will presented here were conducted online and all participants were (monolingual) speakers of Peninsular Spanish. The experimental stimuli consisted of NI presented in contexts of one or two sentences.[1] The verb lexemes were chosen based on token frequency and

structural features. For example, to investigate whether type A prefers realization of the Agent or the Theme, the base had to belong to a verb class that allows the unspecified object alternation. The stimuli were presented along with distractor items in a ratio of 1: 1,5 and were counter-balanced as well as pseudo-randomized. Acceptability was measured on a 7-point Likert-scale. The data were analysed using the Wilcoxon signed-rank tests and a general linear mixed model (Schütze & Sprouse 2013).

Results

Experiment 1 (16 participants / 192 NI) shows that type A, used with a transitive base, prefers realization of the Agent over the Theme, cf. (5) a. vs. b. A clear preference has, however, only been found for generic interpretations like those exemplified in (5) ($p < 0{,}001$).

(5) a. El escribir *del autor* es obsesivo.
'DET writing of the author is obsessive.'
b. El escribir *de un guion* suele ser tedioso.
'DET writing of a screenplay is usually tedious.'

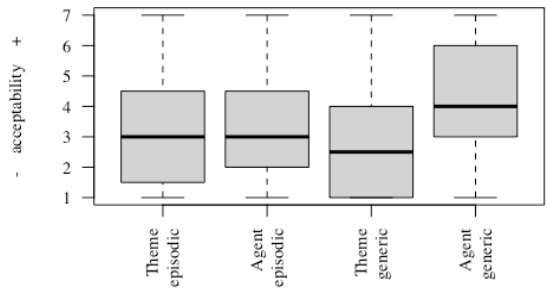


Figure 1. Experiment 1 – Type A (Theme vs. Agent; episodic vs. generic).

Experiment 2 (28 participants / 224 NI) shows that type B is far more acceptable than type A in both episodic and generic readings ($p < 0{,}001$ in each case). Furthermore, the results reveal that type B clearly prefers generic over episodic event interpretations, cf. (6) a. vs. b. ($p < 0{,}01$).

(6) El utilizar productos químicos …
'det using chemical products …'
a. … debería siempre ser el último remedio.
'… should always be the last resort.'
b. … está avanzando por toda la ciudad.
'… is gaining ground throughout the city.'

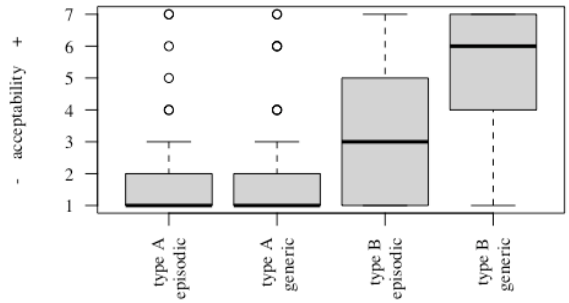


Figure 2. Experiment 2 – Type A vs. B; episodic vs. generic.

Conclusions

The acceptability judgment experiments have revealed fine-grained differences which informally collected data or corpus findings could not have yielded and show that NI can be better analyzed in terms of graded (un)acceptability than in terms of categorical (un)grammaticality. A potential drawback could be seen in the fact that only a limited number of verb lexemes (12 and 8) could be tested with this rather time-consuming method.

It appears that language-specific constraints regarding argument realization interact with possibly more general interpretational preferences of syntactic nominals that differ in their degree of nominalizations. This is indicated, at least, by the preference for generic readings attested for type B.

Notes

1. For reasons of space, the NI are presented here with only their matrix predicate and not the complete context.

References

Demonte, V., Varela S. 1997. Los infinitivos nominales del español. Signo y Seña 4, 125-154.

Featherston, S. 2007. Data in generative grammar: The stick and the carrot. Theoretical Linguistics 33, 269-318.

Iordăchioaia, G., Soare, E. 2015. Deverbal nominalization with the down operator. In Aboh, E.O., Schaeffer, J.C., Sleeman, P. (eds.) 2015. Romance Languages and Linguistic Theory 2013. Selected papers from „Going Romance“ Amsterdam 2013, 223-238, Amsterdam, John Benjamins.

Schütze, C.T., Sprouse, J. 2013. Judgment Data. In Podesva, R.J., Sharma, D. (eds.) 2013, Research Methods in Linguistics, 27-50, Cambridge, CUP.

The effect of verbal short term memory on receptive language abilities in bilingual and monolingual children with Specific Language Impairment

Ioanna Talli, Stavroula Stavrakaki
Department of Italian, Aristotle University of Thessaloniki, Greece

Abstract

The present study investigates whether the bilingual advantage in verbal STM holds for bilingual children with Specific Language Impairment (SLI) and if so whether it affects lexical and syntactic abilities. Sixteen monolingual and 16 bilingual children with SLI, with Greek as L1 and Albanian or Russian as L2 were compared with two control groups of 20 monolingual and 18 bilingual (L1: Greek, L2: Albanian or Russian) TD children in vocabulary, verbal STM and syntactic comprehension. Both clinical groups were worse than mono-TD group. Bi-SLI children were worse than mono-SLI children in vocabulary, in one verbal STM task and in one syntactic comprehension task. The results indicate parallel patterns of deficits in bi- and mono- children with SLI in the domains of STM and syntax.

Key words: Specific Language Impairment, verbal STM, syntax, bilingualism.

Introduction

The term *Specific Language Impairment* (*SLI*) is used to describe the poor oral language development of a child, which cannot be attributed to neurological, sensorimotor, mental or emotional/psychological deficits (Leonard, 1998/2014). Difficulties in language development concern mainly the domain of morpho-syntax (Stavrakaki & van der Lely 2010) and, more specifically, complex structural representations at the levels of syntax and morphology, such as relative clauses (Adani, Forgiarini, Guasti, & van der Lely 2014; Friedmann, Belletti, & Rizzi 2009). In addition, children with SLI show deficits in phonological short-term memory (STM) (Newbury, Bishop & Monaco 2005). Phonological STM is considered a marker for the diagnosis of children with SLI (Archibald & Gathercole 2006; Conti-Ramsden 2003; Lalioti, Stavrakaki, Manouilidou, & Talli, 2016) and an indicator of language skills and syntactic comprehension in typically developing children (Rodrigues & Befi-Lopes 2009).

Recent studies on typically developing bilingual children indicate advantages in the domain of verbal STM (Morales, Calvo, & Bialystok 2013; Soliman 2014), because they are better in executive functions associated with STM, such as task switching (Adi-Japha, Berberich-Artzi, &

Libnawi 2010), which may benefit language acquisition. This study investigates whether the bilingual advantage in verbal STM holds for bilingual children with SLI and if so whether it affects lexical and syntactic abilities. Specifically, our study examines whether there are any differences between monolingual (mono-) and bilingual (bi-) children with SLI on lexical and syntactic skills and on verbal STM.

Methodology

Participants

There were four groups of children in total, two clinical groups and two control groups: The clinical groups included sixteen monolingual (mean age 8.11 years) and 16 bilingual children with SLI (mean age 8.4 years), with Greek as a first language (L1) and Albanian or Russian as a second language. These children were diagnosed with SLI by experienced speech and language therapists working in public or private centres for speech and language therapy in Greece. All these primary school aged children were reported to show language delay or disorders at the preschool age and continued to show persistent language deficits at the time of testing. The two control groups included 20 (Greek-speaking) monolingual (mean age 9.0 years) and 18 bilingual (Greek as L1) typically developing children (TD) (mean age 8.6 years) respectively. These children were matched to the two clinical groups in chronological age and non-verbal IQ (Raven's Coloured Progressive Matrices). The control groups did not differ significantly in chronological age and in Raven percentile from the two clinical groups.

Tasks

Pre-tests included non-verbal IQ assessed with RAVEN Standard Progressive Matrices (Raven 1947) and receptive vocabulary assessed with Peabody Picture Vocabulary Test (Dunn & Dunn 1981; Greek adaptation: Simos, Sideridis, Protopapas, & Mouzaki 2011). Experimental tasks included verbal STM, assessed with three tasks: a) a non-word repetition task (a list of 24 three to six syllables non-words; Talli 2010, adapted in Greek from EVALEC Sprenger-Charolles, Colé, Béchennec, & Kipffer-Piquard 2005), b) a digit span task (forward and backward digit span subtests from Greek WISC-III, Georgas, Paraskevopoulos, Besevegis, & Giannitsas 1997) c) a sentence repetition task (sentence recall subtest from Diagnostic Test of Verbal Intelligence (DVIQ, Stavrakaki & Tsimpli 1999). Experimental tasks also included a task of syntactic comprehension consisted of 36 sentences (24 object and subject relative clauses, 4 with reflexive verbs and 8 with passive voice). The examiner would read the sentence and then the child had to select one of the four pictures that

matched the sentence s/he heard. Percentages of accuracy scores were took into account.

Results

For the different tasks, ANOVAs were conducted with the 4 groups as the between subject factor. When the effect of group was significant, T-tests were conducted to compare performances of mono-SLI and bi-SLI with mono-TD, between bi-SLI and bi-TD and between the two clinical groups.

Both clinical groups had significantly poorer performance compared to the monolingual TD group in all tasks (except for passives and reflexives for mono-SLI). Bi-SLI children had significantly worse performance than mono-SLI children in vocabulary, in one verbal STM task (sentence repetition) and in one syntactic comprehension task (passives). Comparisons between bi-SLI and bi-TD showed that bi-SLI children were significantly worse than bi-TD children in all tasks (see Table 1).

Table 1. Mean performance (and SDs) of groups on experimental tasks (all in percentages, except for the first two).

	Mono-SLI	Bi-SLI	Mono-TD	Bi-TD
Receptive vocabulary (percentile)	24,9 (19,2)	9,9 (12,8)	65,8 (19,8)	39,2 (23,3)
Digit span (raw score)	9,6 (2,8)	8,4 (2)	12,5 (2,5)	10,8 (2)
Non-word repetition	77,6 (12,9)	70,6 (9,7)	91,6 (5,8)	84,2 (8,3)
Sentence repetition	42,9 (2,4)	39,3 (6,9)	44,8 (0,52)	43,8 (1,6)
Relative clauses	67,5 (16,4)	57,5 (14,3)	81,5 (10,4)	84,4 (5,7)
Passives	66,4 (26,5)	45,3 (25,4)	79,4 (17,8)	89,6 (18,3)
Reflexive verbs	81,3 (26,6)	68,8 (19,4)	93,7 (13,8)	98,6 (5,9)

Discussion - conclusions

The abovementioned results indicate that deficits in receptive vocabulary are more evident in the domain of receptive vocabulary in bilingual children with SLI than in monolingual ones. In conclusion, the verbal STM does not appear to be superior in children with bi-SLI or to affect their performance neither on syntax nor on vocabulary. These findings highlight the severely impaired status of verbal STM in SLI independently of the number of languages these children are exposed to.

References

Adani, F., Forgiarini, M., Guasti, M. T., & van der Lely, H. K. 2014. Number dissimilarities facilitate the comprehension of relative clauses in children with

(Grammatical) Specific Language Impairment. Journal of Child Language *41*(4), 811-841.

Adi-Japha, E., Berberich-Artzi, J., & Libnawi, A. 2010. Cognitive flexibility in drawings of bilingual children. Child Development 81, 1356–1366.

Archibald, L., & Gathercole, S. 2006. Nonword repetition: A comparison of tests. Journal of Speech, Language, and Hearing Research 49, 970–983.

Conti-Ramsden, G. 2003. Processing and linguistic markers in young children with specific language impairment. Journal of Speech, Language and Research 46, 1029–1037.

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

Georgas, J., Paraskevopoulos, I. N., Besevegis, E. and Giannitsas, N. D. 1997. *The Hellenic WISC-III*, Athens: Psychometric Laboratory, University of Athens.

Lalioti, M., Stavrakaki, S., Manouilidou C. & Talli, I. (2016). Subject-Verb Agreement and Verbal Short Term Memory: A Perspective from Greek Children with Specific Language Impairment. *First Language*, 36, 279-294.

Leonard, B. L. 1998/2014. *Children with Specific Language Impairment*. Cambridge, MA: MIT Press.

Morales, J., Calvo, A., & Bialystok, E. 2013. Working memory development in monolingual and bilingual children. Journal of experimental child psychology, 114(2), 187-202.

Rodrigues, A., & Befi-Lopes, D. M. 2009. Phonological working memory and its relationship with language development in children. Pró-Fono Revista de Atualização Científica, 21(1), 63-68.

Simos, P., Sideridis, G.D., Protopapas, A. & Mouzaki, A. 2011. Psychometric Characteristics of a Receptive Vocabulary Test for Greek Elementary Students. Assessment for Effective Intervention, 37, 34-50.

Soliman, A. M. 2014. Bilingual advantages of working memory revisited: A latent variable examination. Learning and Individual Differences, 32, 168-177.

Stavrakaki, S., & van der Lely, H. 2010. Production and comprehension of pronouns by Greek children with specific language impairment. British Journal of Developmental Psychology, 28(1), 189-216.

Talli, I. 2010. Linguistic abilities in Developmental Dyslexia and Specific Language Impairment: A comparative and cross-linguistic approach (PhD thesis). Paris Descartes University and Aristotle University of Thessaloniki.

Integrating Incoming Information into Discourse Model in Tunisian Arabic

Marwa Mekni Toujani
Higher Institute of Languages of Tunis, University of Carthage, Tunisia

Abstract

The endeavor of the present study is to gauge the time course of connecting incoming information to information mentioned earlier in the text that are no longer available in Working Memory (WM). Two word-by-word self-paced reading experiments based on the contradiction paradigm were used. The Group Embedded Figure Test (GEFT) was utilized to explore participants' field-dependency. Data from 96 Tunisians reveals that reading times in the inconsistent condition were higher than the consistent condition in the post-critic region when participants were instructed to adopt the protagonist point of view. However, this difference did not reach the significance level when participants were instructed to read for mere comprehension. The results also reveal that reading times for field-dependent participants were higher than their field-independent counterparts. Thus, this study yields evidence for the dominance of local coherence over global coherence.

Key words: Discourse processing, time-course, field-dependency

Introduction

There are two main lines in discourse processing research. The first one is interested in understanding the type of inferences that constitute discourse representations (Stewart, Kidd, & Haigh, 2009). The second line is interested in the time course of integrating incoming information with the unfolding discourse model (ibid). Incoming information can be integrated as soon as it is available or it is integrated at spillover regions as a wrap-up effect (Guzman & Klin, 2000). This study addresses the second line of research and aspires to answer the following questions: (1) does readers' sensitivity to spatial anomaly affect the time-course of integrating incoming information into the unfolding discourse model in Tunisian Arabic (TA)? (2) do field dependency and task demands affect the time-course of integration in TA?

Experimental methodology

The TA version of Stewart et al.'s (2009) sixteen experimental texts were used to explore the phenomenon of time-course of discourse processing in reading. The texts were translated by the researcher, were checked by an expert, and were finally piloted to check their "naturalness", their coherence and the familiarity of the topics. The texts were developed according to the contradiction paradigm (O'Brien & Albrecht, 1992).

The two non-cumulative center word-by word present experiments were run using PsychoPy2 software (Pierce, 2009) on a laptop. Participants were asked to read texts at their normal pace in order to respond to comprehension questions at the end of each passage (experiment 1) and to adopt the protagonist point of view (Experiment 2).

Each text is composed of eight sentences. The first sentence is an introduction mentioning the spatial location. The second, the third and the fourth sentences are filler materials. The fifth sentence is the target sentence. The sixth sentence is the post-target one. The seventh and the eighth sentences are the concluding sentences. There were two versions of each passage. Each passage appeared once in each of the two conditions (consistent or inconsistent condition). Each passage in both experiments was followed by a yes/no comprehension question.

Besides, the GEFT (Witkin, Raskin, Oltman, & Karp, 1971).was utilized to gauge participants field dependency. The GEFT scores ranged from zero to 17 in this study. Participants were categorized as field-dependent (FD) or field-independent (FI) based on the median split option (Raptis, Fidas, & Avouris, 2016). Participants whose scores ranged from zero to five were considered as FD. Participants whose scores ranged from six to 17 were considered as FI.

Ninety-six Tunisian third-year University students of English with native language TA participated in the study (48 participants per experiment).

Results

The total time (in milliseconds) that participants took to read was analyzed for each region. (1) The critical region consists of lexical items that convey the location information. (2) The material following the location items until the end of the target sentence represents the post-critical region. (3) The post-target sentence is the following sentence. Repeated measure ANOVA with both region and condition as within-subjects factors, field dependency as a between-subjects factor and reading times as a dependent variable was used. Both analysis by subjects and analysis by items as random factors were implemented. All analyses reported were significant at the .05 alpha level.

Experiment 1

There was no significant effect of condition in the three regions. However, there was a main effect of field-dependency $F1\ (1, 46) = 9.148$, $p = .004$ with FD participants spent more times reading the texts in both conditions ($M = 2850.242$, $STD = 125.603$) than FI participants did ($M = 2326.513$, $STD = 115.538$). Comprehension question response accuracy was 92%.

Experiment 2

For the post-critical region, there was a significant effect of condition for item analysis F2 (1, 15) = 7.284, p=.01 with this region is more slowly read in the inconsistent condition (M = 2783.62, STD =1604.656) than the consistent condition (M = 2362.69, STD = 1235.421). However, for analysis by participants, the difference between inconsistent and consistent condition only approached significance level F1 (1, 46) = 2.843, p=.09. For both the critical region and the post-target region, there was no effect of condition. There was also no main effect of field-dependency F1<1. Comprehension question response accuracy was 95%.

Discussion and conclusions

The results of the first experiment were in line with McKoon and Ratcliff's (1992) Minimalist Theory. That is, only is information essential for establishing local coherence (connecting the incoming information to one or two prior sentences) is integrated during reading. Global coherence is only maintained online when there is a local coherence break or when readers are instructed to do so (ibid).

Based on the results of the second experiment, integrating incoming information into discourse model is delayed as it is a part of a wrap-up operation. This result is constant with Guzman and Klin's (2000) results. However, this delayed online integration only occurs when participants were instructed to adopt the protagonist point of view. Again, this result endorses the Minimalist Approach.

As the spatial information was the one manipulated in the present experiments, the results are basically in line with Smith and O'Brien's (2012) results. As they argued, participants do not track spatial information if they are not instructed to do so. This instruction triggered their deep processing.

Hence, a revisited version of Cook and Myers (2004) two-stage framework can be suggested. The first stage consists of linking the information at hand with information that is in WM or that is easily accessed from long-term memory (McKoon & Ratcliff, 1992). This stage is automatic and occurs no matter the conditions are. The second stage only occurs when readers have special goals aiming at deep processing to meet their high standards of coherence (Van den Broek, Bohn-Gettler, Kendeou, Carlson, & White, 2011). This stage occurs at wrap-up positions. In future research, protagonist characteristics will be manipulated in order to deepen our understanding of the time-course of discourse processing.

Based on the results of both experiments, it can be noticed that FI participants reacted faster to inconsistent materials than FD ones did, which

supports Hannon and Daneman (2001) who argue that FD people have difficulty in detecting anomalies. More interestingly, the effect of field dependency on the time course of text processing was minimized when participants were instructed to engage in an activity that requires deep processing (i.e. adopting the protagonist point of view). In fact, task demands have a primacy effect over individual differences.

Acknowledgment

The Experiments were conducted in partial fulfillment of the requirements for my PhD degree conducted under the direction of Pr. Mohamed Jabeur to whom go my deepest thanks. I am also grateful to Dr. Tarek Hermessi for helping me reshaping the topic and to Dr. Hafedh Halila for editing the translated materials.

References

Cook, A.E., Myers, J.L. 2004. Processing discourse roles in scripted narratives: The influences of context and world knowledge. Journal of Memory and Language, 50, 268-288.

Guzman, A.E., Klin, C.M. 2000. Maintaining global coherence in reading: The role of sentence boundaries. Memory & Cognition, 28, 722–730.

Hannon, B., Daneman, M. 2001. A new tool for measuring and understanding individual differences in the component processes of reading comprehension. Journal of Educational Psychology, 93, 103e128.

McKoon, G., Ratcliff, R. 1992. Inference during reading. Psychological Review, 99, 440–466.Smith, E.R., O'Brien, E.J. 2012. Tracking spatial information during reading: A cue-based process. Memory & Cognition, 40, 791–801.

O'Brien, E.J., Albrecht, J.E. 1992. Comprehension strategies in the development of a mental model. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 777–784. doi:10.1037/0278-7393.18.4.777

Peirce, J.W. 2009. Generating stimuli for neuroscience using PsychoPy. Front. Neuroinform. 2:10. doi:10.3389/neuro.11.010.2008

Raptis, G., Fidas, C., Avouris, N. 2016. Differences of field dependent/independent gamers on cultural heritage playing: Preliminary findings of an eye-tracking study. In Proceedings of International Conference on Digital Heritage (Euromed2016), Nicosia, Cyprus, 2016. Springer. doi:10.1007/978-3-319-48974- 2_22

Smith, E.R., O'Brien, E.J. 2012. Tracking Spatial Information during Reading: A Cue-Based Process. Memory & Cognition, 40, 791-801.

Stewart, A.J., Kidd, E., Haigh, M. 2009. Early Sensitivity to Discourse-Level Anomalies: Evidence From Self-Paced Reading. Discourse Processes,46:1,46 — 69. DOI: 10.1080/01638530802629091

Van den Broek, P., Bohn-Gettler, C., Kendeou, P., Carlson, S., White, M. J. 2011. When a reader meets a text: The role of standards of coherence in reading comprehension. In M.T. McCrudden, J.P. Magliano, G. Schraw (Eds.), Text relevance and learning from text (pp. 123–140). Greenwich, CT: Information Age

Witkin, H.A., Raskin E., Oltman, P.K., Karp, S.A. (1971). A manual for the Group Embedded Figures Test. Palo Alto, CA: Consulting Psychologists Press.

Effects of part of speech: Primitive or derived from word frequency?

Chengxia Wang, Yi Xu
Dept. of Speech, Hearing and Phonetic Sciences, University College London, UK

Abstract

Part of speech (POS hereafter) is known to affect both duration and F0, such that, nouns are longer and higher in F0 than verbs. In this study we tested the hypothesis that the POS effects are actually a word frequency effect, and that this effect is predictable from information theory. We tested this hypothesis by comparing 44 phonologically matched noun-verb pairs in Mandarin. Results show that there were clear effects of word frequency on duration, but no effects on F0. In contrast, no effects of POS were found on either duration or F0. We conclude that there are no primitive POS effects on duration or F0, but the frequency effect on duration may lead to a weak POS effect given sufficient corpus size.

Key words: part of speech effect, frequency effect, information theory, duration, F0

Introduction

In both recognition and synthesis modeling, part of speech (POS) is taken as an essential input feature during training. But it is unclear why it is so important. Some informal research reports that nouns are longer than verbs (Lightfoot 1970, Coker et al. 1973). However, those studies did not take phrasing and word frequency into account. Word frequency is already known to have an effect on some acoustic parameters（Xu et al. 1973, Aylett et al. 2004, 2006）. So it is possible that POS effects are a result of word frequency. Frequency of occurrence is also important in information theory, as it is directly related to information load (Shannon 2001).

In the present study, we seek to establish a clear link between POS effects and word frequency. Our hypothesis is that the POS effect on duration is derivable from a word frequency effect, such that the latter is much more robust than the former. This is done by comparing pairs of nouns and verbs in Mandarin that are phonetically identical.

Method

Stimuli

44 pairs of sentences were composed, each containing a pair of noun-verb homophones, 12 of which were monosyllabic words and 32 were disyllabic words. Each pair of sentences shared the same total length and position and phrasing of the target words. This design therefore tightly controlled the effects of syllable structure, phrasing and sentence length. The word identity

(hence POS) was represented by their corresponding Chinese characters. The nouns were all made-up first names. There was no word stress or focus in the target words. Frequencies of the verbs were taken from Modern Chinese Frequency Dictionary (Beijing Language and Culture University 1986), while the frequencies of the nouns (names) were calculated from an online name dictionary (online 2016).

Participant

Five female and four male native Mandarin speakers were recorded. They were all college students studying in London and were born and raised in Beijing. The recordings were conducted in a quiet room in Scape Shoreditch, London.

Recording Procedures

All stimulus sentences were presented in Chinese on a computer screen in a random order, with three repetitions each (in separate blocks). In total, 88 (sentences) × 9 (subjects) × 3 (repetitions) = 2376 sentences were recorded.

Categorization of word frequency

To be able to compare word frequency effect directly with POS effect, we transformed the gradient frequency into a categorical variable, using the median frequency as the dividing line. This way, each target word was treated as either high frequency or low frequency.

Results

For the POS effect, verbs are shorter than nouns in terms of mean duration. Two one-way ANOVAs show that POS has a significant effect on duration, $F(1,22) = 5.069$, $p = .035$ for monosyllabic words, and $F(1,62) = 12.691$, $p = 0.001$ for disyllabic words.

For the word frequency effect, low frequency words are longer than high frequency words. However, one-way ANOVAs shows no significant effect on monosyllabic words ($F(1,22) = 1.446$, $p = .242$), but a significant effect on disyllabic words ($F(1,62) = 17.477$, $p < 0.001$).

Because POS and frequency affect duration in the same direction, they are potentially confounded. To avoid confounding, two ANCOVAs were conducted with POS as independent variable, frequency as covariate and duration as dependent variable. The results are displayed in Table 1.

Table 1. ANCOVA results for duration (ms) of monosyllabic and disyllabic words.

Levels	Monosyllabic		Disyllabic	
	POS	frequency	POS	frequency
verb/high frequency	170.7	179	328.7	323
noun/low frequency	198.7	196	367.5	367.6
F value	3.559	0.272	3.429	7.532
p value	0.073	0.608	0.069	0.008

For monosyllabic words, neither POS nor frequency has a significant effect on word duration. A possible cause of the disappearance of POS effect is the small amount of data. But an additional factor is that the division of the words into high and low frequency ones mixed up the syllable structures, so that the frequency comparison was no longer based on minimal pairs of homophones.

For disyllabic words, the POS effect is no longer significant, but the frequency effect is highly significant. This may have been due to several causes: (1) there were more disyllabic than monosyllabic words, (2) some high frequency words had no homophone counterparts with low frequency, and (3) frequency did affect duration significantly. To rule out the second possibility, we looked into each homophone pair of high and low frequency words. If a low frequency noun-verb pair had a complex syllable structure, which may increase duration, it was deleted. After the deletion, 27 pairs were left. Again the frequency effect was significant while the POS effect was not (Table 2).

Table 2. ANCOVA results for duration (ms) of disyllabic words.

Levels	POS	frequency
verb/high frequency	322	319
noun/low frequency	363	364
F value	2.286	5.241
p value	0.137	0.026

Also tones can affect F_0, so sentence pairs in which target words did not have identical tones were then excluded from the data, leaving 12 pairs of sentences with monosyllabic words and 15 pairs with disyllabic words. A set of one-way ANOVAs and two ANCOVAs were performed, neither POS nor frequency effects on F_0 were found.

Discussion and conclusion

This study is a preliminary test of the hypothesis that the effect of part of speech on duration is derived from the effect of word frequency, such that the latter is much more robust than the former. The test was performed by comparing pairs of nouns and verbs in Mandarin that are phonetically identical. Results show that there was no significant POS effect on duration or F_0, but there was a significant frequency effect on duration. Given that POS does show a difference in means between nouns and verbs in the same direction as the frequency effect (Tables 1 and 2), it can be concluded that this weak POS effect is derived from a more primitive effect of frequency.

Such frequency effect is consistent with information theory. That is, speakers are under a general pressure to convey as much information as possible in a given amount of time, and this pressure would lead to each word being assigned as little time as possible. But the pressure is balanced by another pressure of communication, i.e., each word also needs to be pronounced as clearly as necessary, so as not to be misheard. But the chance of being misheard can be reduced by the word's predictability, which can be improved by its frequency. So, words that are of higher frequency can afford to have less time, and thus less full articulation. Verbs, as a group, have higher frequency than nouns, and so are likely to be given less articulation time. But ultimately, it is the frequency of each individual word that partially determines its duration. This seems to be supported from the current data. Given the limited size of the dataset in this study, we are making our conclusion cautiously, however. More research in this direction is needed.

References

Aylett, M., Turk, A. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. Language and speech, 47, 31-56.

Aylett, M., Turk, A. 2006. Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. The Journal of the Acoustical Society of America, 119(5), 3048-3058.

Beijing Language and Culture University. 1986. Modern Chinese Frequency Dictionary.

Coker, C., Umeda, N. and Browman, C. 1973. Automatic synthesis from ordinary English test. IEEE Transactions on Audio and Electroacoustics,21(3), 293-298.

Lightfoot, M. J. 1970. Accent and time in descriptive prosody. Word, 26(1), 47-64.

Names. 2016. [online] Available at: http://www.sosuo.name/tong/.

Shannon, C. E. 2001. A mathematical theory of communication. ACM SIGMOBILE Mobile Computing and Communications Review, 5(1), 3-55.

Xu, Y., Wang, M. 2009. Organizing syllables into groups—Evidence from F0 and duration patterns in Mandarin. Journal of phonetics,37(4), 502-520.

Prosodic focus in three northern Wu dialects: Wuxi, Suzhou and Ningbo

Bei Wang[1], Yuanyuan Zhang[2], Yi Xu[3], Hongwei Ding[2]
[1]Institute of Chinese Minority Languages and Literature, Minzu University of China
[2]Institute of Cross Language Processing Center, Shanghai Jiao Tong University
[3]Department of Speech, Hearing and Phonetic Sciences, University College London

Abstract

The present study investigated three northern Wu dialects: Wuxi, Suzhou, and Ningbo. It is found that, in all three dialects, focus is encoded by increasing the maximum F_0 and duration of focused words, and lowering and compressing the F_0 and pitch range of post-focus words. These results are consistent with previous findings about Wu dialect in Shanghai. Northern Wu dialects therefore seem to encode focus in a similar way to Beijing Mandarin, but different from many languages/dialects spoken in southern China. This finding, together with evidence from gene studies and migration history of Wu areas, provide further support for the *inheritance hypothesis* of PFC, according to which all languages with PFC are descendants of a common proto-language in the Middle East.

Key words: Wu dialect, prosody, focus, post-focus-compression (PFC)

Introduction

Focus is the highlighting of part of a sentence against the rest of the sentence, motivated by discourse context. A focused word is generally realized with increased F_0, duration and intensity. Xu et al. (2012) found that, for the post-focus part, two different patterns have been reported across languages: with post-focus lowering and compression of F_0 and intensity (PFC), or without PFC. PFC has been found that in many languages/dialects spoken in northern and middle China, e.g., Uygur, Ando Tibetan, Lasha Tibetan, Beijing Mandarin, Nanchang dialect, etc., but is absent in many southern languages/dialects, e.g., Taiwanese , Deang, Yi and Wa, etc, as summarized in Xu et al. (2012). They proposed an *inheritance hypothesis*, according to which PFC is inherited from a common proto-language shared by all languages with the feature.

Wu area in China is especially interesting for testing the *inheritance hypothesis* because (1) there are contrary results on the presence of PFC in Wu dialects; (2) the formation of Wu dialects involves influences from both PFC and non-PFC languages in its history. According to Pan (2009), among many others, although Wu dialects are heavily influenced by Kam-Tai languages back to Chunqiu Dynasties (770-476 B.C.), modern Wu dialects (possibly formed in Sunwu era (220-280)) are mainly spoken by people who

immigrated from the north. Wen et al. (2004) showed that the contribution of Y chromosome and mitochondrial DNA (mtDNA) from northern Han is about 80% in the Wu population (Jiangsu and Shanghai). Thus, the *inheritance hypothesis* would predict that Wu languages have PFC just like most northern Mandarin dialects. The current study is an investigation of prosodic focus in the Wu dialects of Wuxi, Suzhou, and Ningbo.

Method

Corrective focus was elicited by a preceding sentence, in a structure as "It is not my brother. My sister poured noodle soup". In each dialect, two declarative sentences were constructed to examine focus in four conditions: initial (IF), medial (MF), final (FF) and neutral (NF).

There were 6 Wuxi speakers (4 F, 2 M, age 24-52), 8 Suzhou speakers (4F, 4M, age 45-63), and 8 Ningbo speakers (5F, 3M, age 24-45).

The recording was digitized into a computer (Lenovo Z475) by 16Bit/44.1k using audio sound (Yamaha Steinberg CI2) and a condenser microphone (SHUER Beta 53). Each speaker recorded all the 8 sentences three times in separately randomized blocks.

The data extraction was done with the Praat script ProsodyPro (Xu, 2013), which generated 10 normalized F0 points, maximum F0, and duration for each manually segmented syllable. The F0 values were converted from Hz to semitones (st) using the following formula:

$$f_{st} = 12 \cdot \log_2(F_0)$$

Results

F_0

Fig. 1 displays the time-normalized F_0 contours of one sentence in each dialect, with four focus conditions overlaid in one figure. The other sentences showed the same pattern.

It can be seen from Fig. 1 that, in all three dialects, focus raises the maximum F_0 of the on-focus word, and compresses the maximum F_0 of the post-focus part as compared to the neutral focus condition, but leaves the pre-focus part largely intact.

Table 1 shows mean maximum F0 of the three target words averaged across the two sentences in the three dialects, broken down by the four focus conditions. Two-way repeated measures ANOVA with word position and focus condition as the independent variables were carried out for each dialect. Focus has significant effects in all the three dialects (Wuxi: $F(3, 15)=6.45^*$, Suzhou: $F(3, 21)=14.73^{**}$, Ningbo: $F(3, 21)=14.31^{***}$).

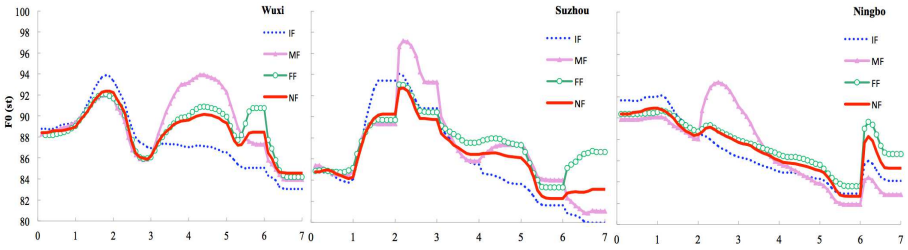


Figure 1. F_0 contours of one sentence in the three Wu dialects.

	Word	IF	MF	FF	NF
Wuxi	W1	94.5	92.6	92.3	92.7
	W2	92.1	92.7	91.5	91.7
	W3	86.6	91.3	90.3	89.2
Suzhou	W1	95.2	94.4	92.5	92.9
	W2	91.1	95.3	92.2	91.3
	W3	83.6	86.8	90.7	87.1
Ningbo	W1	93.9	92.0	92.5	92.6
	W2	87.7	92.6	89.8	89.7
	W3	85.9	85.6	89.8	88.4

Table 1. The Maximum F_0 (st) of the three target words.

Word Duration

Table 2 shows mean duration of the three target words under different focus conditions in each dialect, averaged across three repetitions of the two sentences by all speakers.

	Word	IF	MF	FF	NF
Wuxi	W1	417	375	369	368
	W2	412	492	432	431
	W3	440	444	546	485
Suzhou	W1	528	482	451	468
	W2	390	484	416	416
	W3	436	450	549	495
Ningbo	W1	337	314	304	310
	W2	340	374	366	355
	W3	456	450	496	484

Table 2. The word duration (ms) of the three target words.

It can be seen in Table 2 that on-focus words are lengthened in all three dialects. Pre- and post-focus words do not differ much from the neutral focus words. Two-way repeated measures ANOVAs with focus and word position as two independent variables show that focus has main effect in all the dialects (Wuxi: F(3, 15)=4.7*; Suzhou: F(3,21)=3.46***; Ningbo: F(3, 21)=3.16*).

Discussion and Conclusions

This study investigated focus encoding in Wuxi, Suzhou, and Ningbo. It is found that, in all these dialects, focus not only raises the maximum F0 and lengthens the duration of the on-focus word, but also compresses and lowers the F0 of the post-focus words. Prosodic focus in these Wu dialects thus shows a similar pattern to Shanghai (Chen, 2008; Selkirk & Shen, 1990). Scholtz (2012) has studied a southern Wu dialect (Wenzhou) and showed PFC applied in the initial focus condition. Overall, the results are consistent with the prediction by the *inheritance hypothesis* of PFC. Nevertheless, more languages/dialects need to be studied, especially southern Wu dialects.

Acknowledgements

This paper was jointly supported by the Program for New Century Excellent Talents in University (NCET-12-0584) to B.W and the Interdisciplinary Program of Shanghai Jiao Tong University (14JCZ03) to H-W.D.

References

Pan, W.Y. 2009. The Historical Background of the Formation of Wu Dialects（吴语形成的历史背景）. Dialect（方言), 3, 193-203.

Selkirk, E., Shen, T. 1990. Prosodic domains in Shanghai Chinese. In *The Phonology-Syntax Connection*. S. Inkelas and D. Zec. Chicago: University of Chicago Press pp. 313-37.

Wen, B., Shi, H., Ren, L., Xi, H., Li, K., Zhang, W., . . . Xiao, C. 2004. The origin of Mosuo people as revealed by mtDNA and Y chromosome variation. Science in China Series C: Life Sciences, 47(1), 1-10.

Xu, Y. 2013. ProsodyPro — A tool for large-scale systematic prosody analysis. In *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, Aix-en-Provence, France**:** 7-10.

Xu, Y., Chen, S.-w., Wang, B. 2012. Prosodic focus with and without post-focus compression (PFC): A typological divide within the same language family? The Linguistic Review, 29, 131-147.

Scholtz, F. 2012. Tone sandhi, prosodic phrasing, and focus marking in Wenzhou Chinese. Ph.D. thesis, Leiden University, the Netherlands.

The acquisition of *plus-que-parfait* in French by Chinese learners in French L2 and L3

Xingzi Zhang
Langage et langues, Université Sorbonne Nouvelle, France

Abstract

The purpose of this research is to show the differences of French *plus-que-parfait* acquisition by Chinese learners in French seconde language (L2) and third language (L3) at the initial and intermediate level. Through the comparison of written samples of four groups (native speakers and learners), we found that (1) two groups of French learners sometimes uses the *passé composé* to replace the *plus-que-parfait*; (2) FL2 learners use the subordination with perfective value or lexical means to mark the double anteriority, which is influenced by their L1 Chinese; (3) FL3 learners, who have learned English as L2, are very familiar with the double anteriority indicated by the pluperfect. This likely facilitates the acquisition of the *plus-que-parfait.*

Key words: Chinese learners, *plus-que-parfait,* double anteriority, FL2, FL3

Introduction

The acquisition of French past tense as a foreign language interests many researchers. Many previous studies (Kaplan 1987, Bergström 1995, Kihlstedt 2002, Ayoun 2004) concerning the distinction between *the passé compose* and *imparfait* have been published. The *plus-que-parfait* emerges later among learners (Labeau 2002, Howard 2005, Sun 2006) and receives less attention from researchers. Howard (2005) found that advanced English learners often use the *passé composé* instead of the *plus-que-parfait*, and they often express the double anteriority by lexical or grammatical means. Howard indicates that the use of *plus-que-parfait* is stricter than that of the pluperfect:

1. I had to return home because I forgot / had forgotten my wallet.
   J'ai dû rentrer à la maison parce que j'avais oublié mon portefeuille.
   (Howard 2005: 70)

In this example, one could either use the *preterit* or the *pluperfect* in English to mark the double anteriority, while in French, only the *plus-que-parfait* is acceptable.

In Chinese, there is no equivalent grammatical marques as *pluperfect* in English and *plus-que-parfait* in French. Howard (2005) observed the English learners of French at the advanced level, but at the initial and intermediate

level, how do Chinese learners express the double anteriority in French? This will be explained in the following sections.

Methodology

Written samples of learners who study French as second language (L2) or third language (L3) - English as L2, were compared and the similarities and differences were determined. The written sample is a narrative based on a segment from the silent film “Modern Times”. A total of four groups were included in this study. The two control groups consisted of a group of Chinese natives (CN, n = 8) and a group of French natives (FN, n = 8). The two experimental groups consisted of a group of Chinese FL2 learners (two levels, which were classified by learning time: initial, n = 6 and intermediate, n = 6) and FL3 (also two levels) were included in the study.

Results and Conclusion

Table 2. Use of *plus-que-parfait* of all verbal tenses by Chinese learners and French natives

	FL2 initial	FL3 initial	FL2 intermediate	FL3 intermediate	FN
n	2	0	3	6	8
%	1.3%	0%	1.9%	2.8%	3%

In our corpus, we found that the use of *plus-que-parfait* is limited, not only by learners, but also by French natives. In the initial stage, FL3 learners had not learned *plus-que-parfait* when the data was collected, therefore, it is expected that there is no use of *plus-que-parfait* by FL3 in the initial stage. With the development of acquisition, the use of this tense increased in both groups. The methods used by the learners to compensate the value of *plus-que-parfait* were analyzed and are shown in table 2.

In short, the passé composé is an important method to replace the *plus-que-parfait* for learners of French with different L1s and even for learners who study French as L3. Furthermore, FL2 learners, influenced by their L1, combine the lexical and grammatical methods to replace the *plus-que-parfait* value. It should also be noted that FL3 learners more frequently use the *plus-que-parfait* than FL2 learners in the intermediate stage, although the FL3 learners studied this tense later than FL2 learners. Thus, it is possible that the L2 English facilitate the *plus-que-parfait* learning in L3.

Table 3. Methods to compensate the value of *plus-que-parfait* in both groups of Chinese learners and the ways to indicate double anteriority in Chinese

Initial level (Chinese learners)	Intermediate level (Chinese learners)	Native Chinese
	Passé composé used in relative subordinations by two groups of learners : "c'était… qui" (*It is…that/who*) structure	The structure « shi…de » (It is…that/who)
		Lexical methods : « gānggāng » ou « gāngcái » (just now)
Passé composé used in indirect speech : two groups	Passé composé used in indirect speech: FL3 group	Indirect speech with perfective value (indirect speech+ suffix *-le*)
	Subordination introduced by « avant … » (before): one FL2 learner	Subordination + lexical «前qián… » (avant) (before)
Passé composé lonely: a FL2 learner		No marks
Causal subordination + passé composé/imparfait/ illegible verbal form : two groups	Causal subordination + plus-que-parfait : FL2 learners	

References

Ayoun, D. 2004 The effectiveness of written recasts in the second language acquisition of aspectual distinctions in French : A follow-up study. Modern language journal 88, 31-35.

Bergström, A. 1995. The expression of past temporal reference by English-speakers of French. Phd thesis, The Pennsylvania State University.

Howard, M. 2005. The emergence and use of the plus-que-parfait in advanced French interlanguage. In Focus on French as a foreign Language: Multidisciplinary Approaches, J.-M. Dewaele (éds.), 63-85. Clevedon: Multilingual Matters.

Kaplan, M. 1987. Developmental patterns of past tense acquisition among foreign learners of French, in Van Patten, B. (éds.) Foreign language learning: A research perspective. Rowley MA, Newbury House, 52-59.

Kihlstedt, M. 2002. Reference to past events in dialogue: The acquisition of tense and aspect by advanced learners of French. In Salaberry R, Shirai, Y. (éds). The acquisition of tense-aspect morphology. Amsterdam-New York, John Benjamins, 323-361.

Labeau, E. 2002. The Acquisition of French Past Tenses by Tutored Anglophone Advanced Learners: Is Aspect Enough? Thèse de doctorat, Aston University.

Sun, J.-L 2006. L'acquisition de la temporalité par les appprenants sinophones de français langue étrangère et par les apprenants francophones de chinois langue étrangère. Phd thesis. Université Sorbonne Nouvelle.

Index of names

ExLing 2017
Ebook ISSN: 2529-1092
Ebook ISBN: 978-960-466-162-6