



International Speech Communication Association

ExLing 2011

Proceedings of 4th Tutorial and Research Workshop on
Experimental Linguistics

25-27 May 2011, Paris, France

Edited by Antonis Botinis

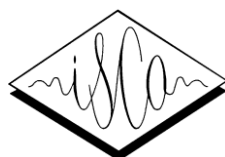


université
PARIS
DIDEROT

Université
Paris Diderot



National and Kapodistrian
University of Athens



International Speech Communication Association

ExLing 2011

Proceedings of 4th Tutorial and Research Workshop on
Experimental Linguistics

25-27 May 2011, Paris, France

Edited by Antonis Botinis



Paris Diderot University



National and kapodistrian
University of Athens

ExLing 2011
Proceedings of 4th Tutorial and Research Workshop on Experimental Linguistics

Ebook ISSN: 2529-1092
Ebook ISBN: 978-960-466-158-9
Copyright © 2011 Antonis Botinis

This volume includes the proceedings of ExLing 2011, the 4th Tutorial and Research Workshop on Experimental Linguistics, in Paris, France, 25-27 May 2011. This is a follow up conference, which is regularly organised from 2006, in Athens, Greece, and thereafter.

Last time in Athens, in summer 2010, we discussed the possibility of organising the next meeting in Paris. Philippe Martin volunteered to take part of the organisation and promote the research paradigm on Experimental Linguistics. In this spirit, we have gathered in Paris 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 a variety of disciplines involved, from phonetics and language typology to psycholinguistics and language acquisition. We are also happy to see young scientists accelerating and adopting innovative ways in the scientific study of language. Most crucially, our initial attempt has gain ground and is becoming an established circle of a new generation of linguists.

As usual, our colleagues are coming from a variety of different parts of the world and we wish them a prosperous exchange of scientific experience and a joyful stay in Paris. This is indeed the core of the ISCA events, which promote new ideas as well as international scientific exchange.

We would like to thank all participants for their contributions and the International Advisory Committee as well as ISCA for its support in organising this conference. We also thank the University of Athens for the publication of the proceedings as well as our students from the University of Paris Diderot for their assistance.

Antonis Botinis

Contents

Keynote lecture

<i>Functions and mechanisms in linguistic research — Lessons from speech prosody</i>	1
Yi Xu	
<i>Qualitatively similar automatic semantic priming in native and non-native speakers</i>	11
Carrie A. Ankerstein	
<i>Structural priming and the phrasal/clausal distinction: The case of concealed questions</i>	15
Gözde Bahadır, Maria Polinsky	
<i>Syntactic recursion and theory-of-mind reasoning in agrammatic aphasia</i>	19
Zoltán Bánréti, Éva Mészáros	
<i>Recursion in language, theory-of-mind inference and arithmetic: aphasia and Alzheimer's disease</i>	23
Zoltán Bánréti, Éva Mészáros, Ildikó Hoffmann, Zita Örley	
<i>Linguistic and non-linguistic investigation of motion events</i>	27
Ayşe Betül Toplu, Deniz Zeyrek	
<i>Semantic priming at the sentence level: causal vs. purposive because</i>	31
Joanna Blochowiak, Gözde Bahadır	
<i>Subject gaps in German coordinative structures – Empirical evidence for a gradient phenomenon</i>	35
Petra-Kristin Bonitz, Anke Holler	
<i>Prosody and quantifier semantics in Greek</i>	39
Antonis Botinis, Aikaterini Bakakou-Orphanou, Anthi Chaida	
<i>Phonology and phonetics of Greek palatalisation</i>	43
Antonis Botinis, Anthi Chaida, Evgenia Magoula	
<i>Perception of French, Belgian and Swiss accents by French and Belgian listeners</i>	47
Philippe Boula de Mareüil, Alice Bardiaux	
<i>Evaluating speech samples designed for the Voice Profile Analysis Scheme for Brazilian Portuguese (BP-VPAS)</i>	51
Zuleica Antonia de Camargo, Sandra Madureira, Luiz Carlos Rusilo	
<i>Dynamic differences in child bilinguals' production of diphthongs</i>	55
Vincent Chanethom	

4 th ISCA Workshop on Experimental Linguistics ExLing 2011	iii
<i>Processing Russian inflectional morphology: A PET study of verb generation</i>	59
Tatiana Chernigovskaya, Kira Gor, Galina Kataeva, Alexander Korotkov, Maxim Kireev, Kristina Memetova, Svyatoslav Medvedev	
<i>The terminal contour of Italian semi-spontaneous instructions</i>	63
Domenico Di Russo	
<i>Models or strategies? On the perception of ambiguous words</i>	67
Anzhelika Dubasova	
<i>Spectral properties of fricatives: a forensic approach</i>	71
Natalie Fecher	
<i>Vocabulary learning strategies among advanced Turkish learners</i>	75
Ioannis Galantomos	
<i>Vowel-colour associations in non-synesthetes: A study with Spanish and Arabic participants</i>	79
Pilar Mompeán Guillamón	
<i>The tip-of-the-tongue phenomenon: Search strategy and resolution during word finding difficulties</i>	83
Nina Jeanette Hofferberth	
<i>L2 Greek morphological agreement</i>	87
Sviatlana Karpava	
<i>Perceptual and lexical priming of syntactic construction in young children</i>	91
Meesook Kim	
<i>Time selected multiple algorithms for reliable Fo tracking in difficult recording conditions</i>	95
Philippe Martin	
<i>Towards a mental representation of vowel height in SSBE speakers</i>	99
Kevin Mendousse	
<i>Prosodic patterns in child speech</i>	103
Roksolana Mykhaylyk	
<i>Tonal and syntactic correlates of focus perception in Greek and Russian</i>	107
Olga Nikolaenkova	
<i>Production of Greek and Turkish vowels by bilingual speakers</i>	111
Elina Nirgianaki, Ougour Chasan, Evgenia Magoula	
<i>The validity of some acoustic measures to predict voice quality settings: trends between acoustic and perceptual correlates of voice quality</i>	115
Luiz Carlos Rusilo, Zuleica Antonia de Camargo, Sandra Madureira	

<i>Lexical stress in Modern Halh Mongolian</i>	119
Yumei Sang	
<i>Pretonic schwa reduction in Dutch: Frequency effects interact with phonotactics</i>	123
Marjoleine Sloos	
<i>Typology and spatial cognition in English, French and Greek: evidence from eye-tracking</i>	127
Efstathia Soroli	
<i>The processing of asymmetric and symmetric sentential conjunction</i>	131
Ellen Thompson, Javier Collado, Maria Omana, Amanda Yousuf-Little	
<i>Identifications of speaker-ethnicity: Attribution accuracy in changeable settings</i>	135
Richard Todd	
<i>Participle Agreement and Clitic Omission</i>	139
Vicenç Torrens	
<i>High functioning autism and prosody of sentence types in Greek: A case study</i>	143
Maria Tripolitou, Anthi Chaida	
<i>Metacommunicative devices in spoken discourse as part of processing distributed cognitive tasks</i>	147
Ilya Utekhin, Tatiana Chernigovskaya	
<i>Perceptual level of intonation in whispered voice</i>	151
Géraldine Vercherand	

Keynote lecture

Functions and mechanisms in linguistic research — Lessons from speech prosody

Yi Xu

Department of Speech, Hearing and Phonetic Sciences, Division of Psychology and
Language Sciences, University College London, UK

Abstract

Human speech, as a collection of complex phenomena, can be explored from many different angles, and interesting data can be generated with various approaches. However, if we aim to substantially improve our understanding of speech, it is advantageous to focus on the communicative functions and the mechanisms that enable these functions. This paper uses prosodic phenomena in speech as examples to illustrate the necessity as well as potential benefits of such function- and mechanism-oriented approaches.

Key words: communicative function, mechanism, speech prosody

Introduction

Human speech, as a collection of complex phenomena, can be studied from many different angles. This is true not only across broad divisions such as syntax, semantics, phonology, etc., but also within each division. In speech prosody, for example, the same set of phenomena can be examined by asking rather different questions: What is the best way to describe melodic events — in terms of level tone or holistic contour (Bolinger, 1986; Cruttenden, 1997; Ladd, 2008; Pierrehumbert, 1980)? How are prosodic patterns related to syntax (Beckman, 1996; Selkirk, 1984)? What are the prosodic correlates of prominence and boundary tones (Hermes & Rump, 1994; Kochanski et al., 2005; Terken, 1991)? What is the essence of speech rhythm (Loukina et al., 2011; Ramu et al., 1999)? Is compression or truncation the chief strategy of tonal variation in a language (Grabe, 1998)? Should tonal alignment be part of the basic description of a language (Ladd et al., 1999; Kohler, 2005)? While these questions all seem very interesting, there is another question that is rarely asked, namely, is answering questions like these the best way to improve our understanding of speech in general and speech prosody in particular?

What most of these questions have in common is that they are mainly about the form of prosody. That is, they address issues that are directly related to the directly observable properties of prosody, while the communicative meanings associated with these properties are left vague. A good example is the Autosegmental-Metrical (AM) model of intonation, in which the intonation components like pitch accent, phrase tone and boundary

tone are all defined in form, while their functional definitions are relatively vague (Ladd, 2008; Pierrehumbert, 1980), and this is despite deliberate efforts to identify the possible functional meanings associated with the proposed categories (Pierrehumbert & Hirschberg, 1990).

An alternative strategy that has been steadily gaining recognition is the function-oriented approaches or functional approaches for short (Hirst, 2005; Kohler, 2005; Xu, 2005). Nevertheless, non-functional approaches still dominate prosody research. Likewise, in many other areas of linguistics, non-functional approaches also dominate. In the rest of this paper, a case will be made for approaches that focus directly on communicative functions and mechanisms that enable their realization.

The case for functional approaches

Of the many reasons for taking functional approaches seriously, the most fundamental is based on consideration of the nature of speech. This can be illustrated by a comparison between speaking and singing, both of which use the human vocal apparatus to control the melody of vocalization. In singing, every note needs to be as accurate as ± 1 semitone from the designate frequency (Dalla Bella et al., 2007; Pfordresher et al., 2010), otherwise the singer will sound out of tune. This is the case because music is form-driven (Patel, 2008), and so perfection in form is of utmost importance. In speech, in contrast, even in a tone language where every tone needs to be quite accurate, the precise F_0 value of any tone varies extensively from speaker to speaker, and from context to context, even among speakers of the same gender and age. But the tones still sound natural and function perfectly, because within individual and each context, different tones are sufficiently dissimilar to each other so that listeners have little difficulty telling them apart, at least in non-adverse conditions (Gandour, 1983). But this is not surprising given that the function of lexical tone is to *distinguish* words/morphemes that are otherwise identical to each other. Thus on the one hand only distinct tonal contrasts would have emerged and survived in a language, on the other hand, serviceable contrasts do not require precisions as high as in music. Such a natural design makes it possible for individuals who cannot sing in tune (10-15% of the population: Dalla Bella et al., 2007; Pfordresher et al. 2010) or detect out-of-tune singing (about 4% of the population: Kalmus & Fry, 1980; Nan et al., 2010) to still function normally in speech communication. The nature of speech therefore determines that speech is not about forms that are analogous to musical notes, but about the functions behind the forms, and so it is function-driven (Patel, 2008).

Lessons from phonemes

Functions in speech, however, are often elusive. Is it really possible to make a rigorous pursuit of functions? The answer is definitely yes, because functional approaches are not strangers to us. One of the most significant progress in linguistics, i.e., the development of the notion of phoneme, has actually established a basic principle of functional approaches. Based on this principle, a phonemic contrast is established, first and foremost, based on its ability to distinguish words or grammatical functions. Phonetic variants that do not distinguish one word from another, e.g., the many variety of /r/ in English, are considered as allophones rather than different phonemes. Thus lexical contrast is the *defining property* rather than *subordinate* or *accompanying property* of phonemes.

Another important aspect of the phonemic notion, which is not frequently highlighted, is that phonemes are not meaning carriers themselves, but serve only to distinguish lexical items from one another. The specific lexical meanings are morphologically rather than phonologically defined, and only occasionally, e.g. in the case of onomatopoeia, is there any direct link between phonetic form and lexical meaning. This is very different from the notion of “intonational meaning”, according to which components of intonation are directly meaningful (cf. Ladd, 2008 for a review).

Following the phonemic tradition, then, the differences in form that matter phonologically would be those that make functional contrasts, rather than any observable differences, even if the observation is experimental. But if this is the case, there needs to be serious rethinking about many prosodic patterns that have been deemed “phonological” or important. These may include, e.g., *pitch accent*, *phrase accent* and *boundary tone* in the AM theory of intonation (Gussenhoven, 2004; Ladd, 2008; Pierrehumbert, 1980), patterns that are considered to be important for the description and teaching of intonation, e.g., *nucleus*, *head*, *pre-head* and *tail* in the British nuclear tradition (Halliday, 1967; Kingdon, 1958; O'Connor & Arnold, 1961; Palmer, 1922), rhythmic patterns such as *stress-timing*, *syllable-timing* and *mora-timing* (Ramu et al., 1999; Port, 2003; Warner & Arai, 2001), *prosodic hierarchy* in terms of *prominence* (Beckman, 1996; Selkirk, 1984), etc. In each of these cases, the critical question one may want to ask is, are the proposed categories primarily defined by the functions they serve, or is the functional aspect of the categories treated only as secondary?

Functional ≠ categorical

An important bias that is likely introduced by the phonemic tradition is the primacy given to categorical contrasts. The issue is not only about the laboratory phenomenon of categorical perception, which has already been met with questions (Fujisaki & Kawashima, 1971; Schouten et al., 2003), but also the notion that all linguistic or communicational contrasts have to be

categorical (Ladd, 2008; Liberman & Pierrehumbert, 1984). What needs to be recognized is that probably due to fluid vocabulary changes as well as phonological crowding (as exemplified by the rapid abstraction and categorization of newly invented sign languages over just a few generations: Pietrandrea, 2002), it is difficult to maintain iconicity in lexical construction, especially that involving gradience. For instance, it is hard to imagine representation of object size with a gradient change in vowel colour. For non-lexical functions, however, it is possible to represent gradient contrasts. For example, there is much evidence, though further research is still needed, that the demarcation/grouping function, as marked by F_0 height and domain-final lengthening, is gradient rather than categorical (Wagner, 2005; Byrd & Krivokapić, 2006). Therefore, categoricalness per se probably should not be used as the benchmark for determining whether a contrast is functional.

The need to establish mechanisms

Equally important as the need to focus on functions is the necessity to pursue the underlying mechanisms that encode the functions. This is because without establishing plausible mechanisms, the link between observed forms and underlying functions remains incomplete. Establishing a continuous link is not easy, however. In the case of prosody, for example, at least three degrees of separation between surface prosodic forms and the communicative functions they encode can be identified, namely, *articulatory constraints*, *target reassignment* and *parallel encoding* (Xu, 2004a). Each degree of separation actually involves a different set of mechanism. At the articulatory level are mechanisms like *target approximation*, *target-syllable synchronization*, *cross-boundary state transfer*, *anticipatory dissimilation*, *post-L bouncing*, *vowel intrinsic F_0* , *consonantal perturbation*, etc. (see Xu, 2006 for a review). Because of these mechanisms, directly observed surface acoustic forms can *never* fully resemble the underlying phonetic targets used to encode the functions.

At the level of target assignment, where an underlying pattern is determined for each syllable, there are often language-specific rules that change the targets depending on factors like phonetic context and communicative functions. The most striking example is the tone sandhi phenomenon in many tone languages (Chen, 2000). In Mandarin, for instance, the Low tone has a rising tail when produced in isolation and sometimes sentence-finally. But when it is followed by any other tone the rising tail is missing, and its absence has no plausible articulatory explanations (Xu, 2004b). Also, the Low tone changes into a Rising tone when followed by another Low tone, and again there are no plausible articulatory explanations. More recently, it is shown that target reassignment happens also in English intonation, where the underlying pitch target

associated with a stressed syllable varies across high, fall and rise depending on its position in word and sentence, whether it is in focus, and whether the sentence is a statement or question (Liu & Xu, 2007; Xu & Xu, 2005). Target reassignment, whenever it occurs, makes it difficult to directly recognize functionally relevant targets. But the recognition is not impossible if we employ methods that are sufficiently sensitive.

Beyond articulatory constraints and target assignment, there are also mechanisms that make it possible for multiple functions to be simultaneously encoded. Such parallel encoding of multiple functions, however, further obscures the link between any specific functions and the directly observable prosodic forms. To reveal the actual link for each function, it is necessary to conduct controlled experiments (e.g., Cooper et al., 1985; Eady & Cooper, 1986; Eady et al., 1986; Liu & Xu, 2007; Pell, 2001; Wagner, 2005; Xu, 1999; Xu & Xu, 2005). The observation of such multi-dimensional coding is what has inspired the parallel encoding and target approximation model (PENTA) (Xu, 2005).

Needless to say, any proposed mechanisms, including the ones just mentioned, are open for debate, as scientific hypotheses should be. What is important is to recognize the need to actively pursue them rather than being easily satisfied by seeming descriptive adequacy. Also, just as importantly, once recognized, mechanisms should always be taken into consideration whenever they may apply rather than being used only when convenient.

Potential benefit of function- and mechanism oriented approaches

Functions and mechanisms are not only of theoretical interest or beneficial only within a particular area of research, but they may also have broader impacts. In the case of prosody, for example, better understanding of functions and mechanisms may make the research findings more relevant for language teaching, speech technology, the relation between prosody and other levels of speech and the relation between linguistics and other disciplines. In regard to language teaching, the nuclear tone tradition was actually first developed as a tool for teaching English intonation (Palmer, 1922). However, after being promoted by generations of authors (e.g., Brazil, 1980; Cruttenden, 1997; Crystal, 1969; Halliday, 1967; Kingdon, 1958; O'Connor & Arnold, 1961; Palmer, 1922), its effectiveness in teaching English intonation is still yet to be demonstrated (Atoye, 2005). In this tradition, priority is given to detailed description of melodic contours of intonation in terms of nucleus, head, pre-head and tail. Assuming that the descriptions are reasonably accurate, why aren't they helpful for the students? Given that the communicative functions of these components are not clearly defined (other than that the nucleus is partially equivalent to focus), it would be difficult for students to learn when to use which pattern.

More importantly, not functionally-defined also means that each described melodic pattern may carry multiple functions, due to parallel encoding as discussed above. Teaching such confounded patterns is unlikely to facilitate the learning of true regularities in intonation.

In speech technology, it has been long desirable to significantly improve the prosody of synthesis, with the goal to generate highly natural as well as expressive speech. But the success has been limited so far. From a functional and mechanistic point of view, this is not surprising, because naturalness and expressiveness are not abstract properties. Truly natural and expressive speech, by definition, has to convey rich and appropriate communicative meanings, and should do so in a manner that resembles the human articulation process. Thus improved understanding of functions and mechanisms may therefore help to move speech technology forward toward being able to produce synthetic speech that is truly expressive and natural. Some efforts in this direction have already been made (Bailly & Holm, 2005; Prom-on et al., 2009).

The mechanistic-functional view of speech may also improve our understanding of the link between different levels of linguistic processing. For example, a long-standing issue in prosody is how prosodic structures are linked to syntactic structures. Some accounts favour close links (Chomsky & Halle, 1968; Selkirk, 1984), while others favour relative independence (Beckman, 1996; Shattuck-Hufnagel & Turk, 1996). From a functional perspective, however, the issue can be viewed in an entirely different light. That is, because speech is about exchanging information, both syntax and prosody should be for that purpose. But for the sake of information transmission, if a function is already syntactically coded, there is no need to also encode it prosodically, and vice versa, unless of course, *redundancy* of coding is favoured (Assmann & Summerfield, 2004). Redundant coding is indeed found for various functions. For example, focus is often marked by both syntactic and prosodic means (Féry et al., 2010). Likewise, boundary or grouping information is also likely encoded by both syntactic/semantic and prosodic means (Wagner, 2005). On the other hand, the specifics of such redundancy do not have to be universal. For example, recent research shows that post-focus compression is a highly effective prosodic cue of focus (Chen et al., 2009), but many languages do not have post-focus compression (Xu, 2011; Zerbian et al., 2010). This new finding has now led to hypotheses that may link cross-linguistic distributions of post-focus compression to recent findings in population genetics (Xu, 2011).

Conclusion

This paper has presented arguments for adopting a new perspective in linguistic research which gives high priority to the pursuit of communicative functions and the underlying mechanisms. It is argued that many previously reported phenomena may be viewed in a very different light in this perspective, which may actually enable us to quickly hone in on issues that really matter in speech. In addition to theoretical advantages there are also potential benefits of adopting function-mechanism-oriented approaches.

References

- Assmann, P.F., Summerfield, A.Q. 2004. The perception of speech under adverse conditions, in: S. Greenberg, W.A. Ainsworth, A.N. Popper, R. Fay (Eds.), *Speech Processing in the Auditory System*. Springer-Verlag, New York.
- Atoye, R. O. 2005. Non-native perception and interpretation of English intonation. *Nordic Journal of African Studies* 14, 26-42.
- Bailly, G., Holm, B. 2005. SFC: a trainable prosodic model. *Speech Communication* 46, 348-364.
- Beckman, M.E. 1996. The parsing of prosody. *Language and Cognitive Processes* 11, 17-67.
- Bolinger, D. 1986. *Intonation and its parts: melody in spoken English*. Stanford University Press, Palo Alto.
- Brazil, D.M. Coulthard, M. and Johns, C., 1980. *Discourse Intonation and Language Teaching*. Longman, London.
- Byrd, D., Krivokapić, J., Lee, S. 2006. How far, how long: On the temporal scope of phrase boundary effects. *Journal of the Acoustical Society of America* 120, 1589-1599.
- Chen, M.Y. 2000. *Tone Sandhi: Patterns across Chinese Dialects*. Cambridge University Press, Cambridge, UK.
- Chen, S.-w., Wang, B., Xu, Y. 2009. Closely related languages, different ways of realizing focus. *Proceedings of Interspeech 2009*, Brighton, UK: 1007-1010.
- Chomsky, N., Halle, M. 1968. *The Sound Pattern of English*. Harper & Row, New York.
- Cooper, W.E., Eady, S. J., Mueller, P.R. 1985. Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America* 77, 2142-2156.
- Cruttenden, A. 1997. *Intonation*. Cambridge University Press, Cambridge.
- Crystal, D. 1969. *Prosodic Systems and Intonation in English*. Cambridge University Press, London.
- Dalla Bella, S., Giguère, J.-F., Peretz, I. 2007. Singing proficiency in the general population. *Journal of the Acoustical Society of America* 121, 1182-1189.
- Eady, S. J., Cooper, W. E. 1986. Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America* 80, 402-416.
- Eady, S.J., Cooper, W.E., Klouda, G.V., Mueller, P.R., Lotts, D.W. 1986. Acoustic characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments. *Language and Speech* 29, 233-251.

- Féry, C., Skopeteas, S., Hornig, R. 2010. Cross-linguistic comparison of prosody, syntax and information structure in a production experiment on localising expressions. *Transactions of the Philological Society* 108(3), 329-351.
- Fujisaki, H., Kawashima, T. 1971. A model of the mechanisms for speech perception: Quantitative analysis of categorical effects in discrimination. *Annual Report of the Engineering Research Institute (Faculty of Engineering, University of Tokyo)* 30, 59-68.
- Gandour, J. 1983. Tone perception in Far Eastern languages. *Journal of Phonetics* 11, 149-175.
- Grabe, E. 1998. Pitch accent realization in English and German. *Journal of Phonetics* 26, 129-143.
- Gussenhoven, C. 2004. *The Phonology of Tone and Intonation*. Cambridge University Press.
- Halliday, M.A.K. 1967. *Intonation and Grammar in British English*. Mouton, The Hague.
- Hermes, D.J., Rump, H.H. 1994. Perception of prominence in speech intonation induced by rising and falling pitch movements. *Journal of the Acoustical Society of America* 96, 83-92.
- Hirst, D.J. 2005. Form and function in the representation of speech prosody. *Speech Communication* 46, 334-347.
- Kalmus, H., Fry, D.B. 1980. On tune deafness (dysmelodia): Frequency, development, genetics and musical background. *Annals of the Human Genetics* 43, 369-382.
- Kingdon, R. 1958. *The groundwork of English intonation*. Longman, London.
- Kochanski, G., Grabe, E., Coleman, J., Rosner, B. 2005. Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustical Society of America* 118, 1038-1054.
- Kohler, K. 2005. Timing and Communicative Functions of Pitch Contours. *Phonetica* 62, 88-105.
- Ladd, D.R., Faulkner, D., Faulkner, H., Schepman, A. 1999. Constant “segmental anchoring” of F0 movements under changes in speech rate. *Journal of the Acoustical Society of America* 106, 1543-1554.
- Liberman, M., Pierrehumbert, J. 1984. Intonational invariance under changes in pitch range and length, in: M. Aronoff and R. Oehrle (Eds.), *Language Sound Structure*. M.I.T. Press, Cambridge, Massachusetts: 157-233.
- Liu, F., Xu, Y. 2007. Question intonation as affected by word stress and focus in English. *Proceedings of The 16th International Congress of Phonetic Sciences, Saarbrücken*: 1189-1192.
- Loukina, A., Kochanski, G., Rosner, B., Keane, E., Shih, C. 2011. Rhythm measures and dimensions of durational variation in speech. *Journal of the Acoustical Society of America* 129, 3258-3270.
- Nan, Y., Sun, Y., Peretz, I. 2010. Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. *Brain* 133, 2635-2642.
- O'Connor, J.D., Arnold, G.F. 1961. *Intonation of Colloquial English*. Longmans, London.
- Palmer, H.E. 1922. *English Intonation, with Systematic Exercises*. Heffer, Cambridge.

- Patel, A.D. 2008. *Music, Language, and the Brain*. Oxford University Press, New York.
- Pell, M.D. 2001. Influence of emotion and focus on prosody in matched statements and questions. *Journal of the Acoustical Society of America* 109, 1668-1680.
- Pfordresher, P.Q., Brown, S., Meier, K., Belyk, M., Liotti, M., 2010. Imprecise singing is widespread. *Journal of the Acoustical Society of America* 128, 2182-2190.
- Pierrehumbert, J., Hirschberg, J., 1990. The meaning of intonational contours in the interpretation of discourse, in: P. R. Cohen, J. Morgan and M. E. Pollack (Eds.), *Intentions in Communication*. MIT Press, Cambridge, Massachusetts: 271-311.
- Pierrehumbert, J. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. dissertation, MIT, Cambridge, MA. [Published in 1987 by Indiana University Linguistics Club, Bloomington].
- Pietrandrea, P. 2002. Iconicity and Arbitrariness in Italian Sign Language. *Sign Language Studies* 2, 296-321.
- Port, R.F. 2003. Meter and speech. *Journal of Phonetics* 31, 599-611.
- Prom-on, S., Xu, Y., Thipakorn, B., 2009. Modeling tone and intonation in Mandarin and English as a process of target approximation. *Journal of the Acoustical Society of America* 125, 405-424.
- Ramus, F., Nespor, M., Mehler, J. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265-292.
- Rump, H. H., Collier, R. 1996. Focus conditions and the prominence of pitch-accented syllables. *Language and Speech* 39, 1-17.
- Schafer, A. J., Speer, S. R., Warren, P., White, D. 2000. Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research* 29, 169-182.
- Schouten, B., Gerrits, E., Hessen, A. v. 2003. The end of categorical perception as we know it. *Speech Communication* 41, 71-80.
- Selkirk, E. 1984. *Phonology and syntax: the relation between sound and structure*. MIT Press, Cambridge, Mass.
- Shattuck-Hufnagel, S., Turk, A.E. 1996. A Prosody Tutorial for Investigators of Auditory Sentence Processing. *Journal of Psycholinguistic Research* 25(2), 193-247.
- Shen, X.S. 1993. The use of prosody in disambiguation in Mandarin. *Phonetica* 50, 261-271.
- Terken, J. 1991. Fundamental frequency and perceived prominence of accented syllables. *Journal of the Acoustical Society of America* 89, 1768-1776.
- Wagner, M. 2005. *Prosody and Recursion*. Ph.D. Dissertation, Massachusetts Institute of Technology.
- Warner, N., Arai, T. 2001. Japanese Mora-Timing: A Review. *Phonetica* 58, 1-25.
- Xu, Y., Xu, C. X. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33, 159-197.
- Xu, Y. 1999. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics* 27, 55-105.
- Xu, Y. 2004a. Separation of functional components of tone and intonation from observed F0 patterns, in: G. Fant, H. Fujisaki, J. Cao, Y. Xu (Eds.), *From Traditional Phonology to Modern Speech Processing: Festschrift for Professor Wu*

- Zongji's 95th Birthday. Foreign Language Teaching and Research Press, Beijing: 483-505.
- Xu, Y. 2004b. Understanding tone from the perspective of production and perception. *Language and Linguistics* 5, 757-797.
- Xu, Y. 2005. Speech melody as articulatorily implemented communicative functions. *Speech Communication* 46, 220-251.
- Xu, Y. 2006. Principles of tone research. *Proceedings of Second International Symposium on Tonal Aspects of Languages, La Rochelle, France*: 3-13.
- Xu, Y. 2011. Post-focus compression: Cross-linguistic distribution and historical origin. *Proceedings of The 17th International Congress of Phonetic Sciences, Hong Kong*.
- Zerbian, S., Genzel, S., Kügler, F., 2010. Experimental work on prosodically-marked information structure in selected African languages (Afroasiatic and Niger-Congo). *Proceedings of Speech Prosody 2010, Chicago*: 100976:1-4.

Qualitatively similar automatic semantic priming in native and non-native speakers

Carrie A. Ankerstein

Department of English, Saarland University, Germany

Abstract

Qualitative and quantitative differences in semantic priming for native and non-native speaker groups were explored using a primed lexical decision task. In addition to response time data, coefficient of variance (CV) for response times was used to investigate quantitative differences in lexical processing between groups. Segalowitz and Segalowitz (1993) argued that differences in CV indicate processing differences. For example, lexical access via automatic routes results in lower CVs, and lexical access via attentional routes results in higher CVs. The current study replicates findings for automatic semantic priming in non-native speakers and contributes behavioural data for the argument that priming in a non-native speaker group can be qualitatively similar to that in native speakers.

Key words: priming, second language acquisition, coefficient of variance, automaticity

Introduction

Semantic priming and lexical access in non-native speakers is widely researched with one of the recent questions being whether priming and lexical access in non-native speakers is qualitatively different to that in native speakers (Frenck-Mestre & Prince, 1997; Phillips, Segalowitz, O'Brien & Yamasaki, 2004). Lexical access is often tested using response timed lexical decision tasks. However, Segalowitz and Segalowitz (1993) pointed out that response time data alone cannot indicate whether a participant is using automatic (direct access) or attentional (e.g., translation or rule checking) processes. In their influential paper, they argued that the coefficient of variance for response times (standard deviation divided by mean response time) can be used to distinguish between automatic (lower CVs) and attentional (higher CVs) processing whilst correcting for differences in response speed.

Most studies using coefficient of variance analysis have focussed on longitudinal changes in proficiency within subjects following training (for a review, see Hulstijn et al., 2009). The question of within subject change is interesting, but so far neglected in this recent wave of research is whether non-native speakers of a language can attain a native-like state. The current study investigated potential differences in semantic priming in native and non-native speakers of English using a lexical decision task. Semantic priming refers to the facilitation of responses to a stimulus as a result of a previous semantically related stimulus. Automatic semantic priming is often

attributed to spreading activation throughout representations in the semantic network (e.g., Masson, 1995).

Word frequency was also manipulated in the experiment. It has been suggested that even native speakers process high and low frequency words differently (Balota, 1994). It has also been proposed that priming effects for low frequency word targets should be stronger than for high frequency word targets (Kinoshita, 1995). This is because response times for high frequency words are already fast and less affected by the “boost” from a prime word. If different processes are indeed involved in the processing of high and low frequency words, then CVs should differ as a function of frequency in native speakers at least.

Methodology

Thirty native speakers of British English (21 females, 9 males) aged between 18–30 years (mean age = 21.2 years) were recruited from the University of Sheffield, England and 24 non-native speakers of English (18 females, 6 males) aged between 18–33 years (mean age = 22.6 years) were recruited from Saarland University, Germany. All non-native speakers had learned English in school from the age of 10 years and were studying English or used English at university.

Stimuli consisted of 90 English prime-target word pairs which appeared in related (spring-season), unrelated (tongue-season) and neutral (BLANK-season) conditions in three counterbalanced lists. Half of the targets were high frequency (rankings from 509-4231) and half were low frequency (rankings from 2-66). Frequency ratings were taken from the Thorndike-Lorge written frequency ratings (Fearnley, 1997). There was a total of 360 stimulus items including 180 prime and target words, 120 phonotactically legal nonwords and 60 word and neutral word (BLANK) filler items. Participants were presented with a single stimulus and asked to judge whether the stimulus was a word or nonword as quickly and as accurately as possible. Participants responded via mouse button press.

Results

Data from one non-native speaker participant was not recorded due to a computer error; data are reported for the twenty-three remaining participants. Correct responses were analyzed and response times of 2000 milliseconds (the time-out value) were eliminated from analysis. Response times \pm 3 SD of the subject's mean were replaced with the subject's mean response time for that condition. Response times (RT) and coefficient of variance for response times (CV) were entered into separate repeated measures ANOVAs with Condition (related, unrelated, neutral) and Frequency (high, low) as

within subjects variables and Group (native, non-native) as a between subjects variable.

For the RT data, there was a significant effect of Frequency ($F(1, 51) = 248.811$, $p < 0.0005$), indicating RTs for high frequency targets were significantly faster than for low frequency targets. There was a significant effect of Condition ($F(2, 102) = 16.300$, $p < 0.0005$). Post hoc tests showed that this was due to the significantly faster RTs for the related condition as compared to the unrelated condition ($t(52) = 3.629$, $p < 0.01$) and the neutral condition ($t(52) = 5.499$, $p < 0.0005$). There were no other significant comparisons, all $p > 0.02$; Bonferroni adjusted p value = 0.02. There was a significant interaction between Condition and Frequency ($F(2, 102) = 6.502$, $p < 0.01$). Post hoc tests showed that RTs for the high frequency related condition were significantly faster than for the high frequency neutral condition ($t(52) = 3.155$, $p < 0.008$) and RTs for the high frequency unrelated condition were faster than for the high frequency neutral condition ($t(52) = 3.185$, $p < 0.008$); Bonferroni adjusted p value = 0.008. There was no significant difference for the high frequency related and unrelated comparison ($t(52) = 0.418$, $p = 0.678$). RTs for the low frequency related condition were significantly faster than the low frequency unrelated condition ($t(52) = 4.202$, $p < 0.0005$) and the low frequency neutral condition ($t(52) = 4.904$, $p < 0.0005$), indicating a significant priming effect for low frequency targets. There was no significant difference for RTs in the low frequency unrelated and neutral comparison ($t(52) = 0.602$, $p = 0.550$). There was a significant main effect of Group ($F(1, 51) = 15.669$, $p < 0.0005$), indicating that RTs for native speakers (mean = 552.1msec) were significantly faster than for non-native speakers (mean = 640.4msec).

For the CV data, there was a significant effect of Frequency ($F(1, 51) = 24.383$, $p < 0.0005$), indicating that CVs for high frequency targets (mean = 0.21) were significantly lower than for low frequency targets (mean = 0.26). There was no significant effect of Condition ($F(2, 102) = 0.414$, $p = 0.662$), indicating that CVs across the conditions were similar. There was a significant interaction between Condition and Frequency ($F(2, 102) = 4.941$, $p < 0.01$), indicating different CV values for high and low frequency targets across conditions. Post hoc tests showed that CVs for the high frequency unrelated and neutral conditions were significantly different ($t(52) = 3.706$, $p < 0.008$). There were no other significant differences, all $p > 0.008$; Bonferroni adjusted p value = 0.008. There was no significant main effect of Group ($F(1, 51) = 0.164$, $p = 0.688$), indicating that native (mean = 0.23) and non-native speaker (mean = 0.24) CVs were similar.

Discussion

The results replicate previous findings of semantic priming in a second language, but in contrast to previous studies (e.g., Frenck-Mestre and Prince, 1997), the current study used coefficient of variance of response times, in addition to response time data, to explore possible qualitative semantic priming differences between native and non-native speakers.

The results yielded a number of important findings. Firstly, the current data provide empirical behavioural evidence for the difference in processing and priming of high and low frequency words in native (and non-native) speakers using the coefficient of variance for response times. CVs in both groups were lower for high frequency words indicating more automatic processing of these words. Priming for low frequency word targets was also stronger than for high frequency word targets in both groups. Secondly, the study replicates automatic semantic priming in a second language. Finally, the study shows empirical support in the form of behavioural data, notably the coefficient of variance, for the argument that semantic priming in non-native speakers can be qualitatively similar to priming in native speakers, even though non-native speaker may have longer response times than native speakers.

References

- Balota, D.A. 1994. Visual word recognition: the journey from features to meaning. In Gernsbacher, M.A. (ed.) 1994, *Handbook of Psycholinguistics*, 303-358. San Diego, CA: Academic Press.
- Fearnley, S. 1997. MRC Psycholinguistic database search program. *Behavior Research Methods, Instruments & Computers* 29(2), 291-295.
- Frenck-Mestre, C., Prince, P. (1997). Second language autonomy. *Journal of Memory and Language* 37, 481-501.
- Hulstijn, J.H., Van Gelderen, A., Schoonen, R. 2009. Automatization in second language acquisition: what does the coefficient of variation tell us? *Applied Psycholinguistics* 30, 555-582.
- Kinoshita, S. 1995. The word frequency effect in recognition memory versus repetition priming. *Memory & Cognition* 23, 569-580.
- Masson, M.E.J. 1995. A distributed-memory model of semantic priming. *Journal of Experimental Psychology-Learning, Memory and Cognition* 21(1), 3-23.
- Phillips, N.A., Segalowitz, N., O'Brien, I., Yamasaki, N. 2004. Semantic priming in a first and second language: evidence from reaction time variability and event-related brain potentials. *Journal of Neurolinguistics* 17, 237-262.
- Segalowitz, N., Segalowitz, S.J. 1993. Skilled performance, practice, and the differentiation of speed-up from automatization effects: evidence from second language word recognition. *Applied Psycholinguistics* 14, 369-385.

Structural priming and the phrasal/clausal distinction: The case of concealed questions

Gözde Bahadır¹, Maria Polinsky²

¹Cognitive Science, Middle East Technical University, Turkey

²Linguistics, Harvard University, USA

Abstract

This paper investigates whether structural priming is sensitive to the phrasal vs. clausal nature of the constructions it tests. To that end, we examine NPs that receive a question-like interpretation when embedded under certain predicates. These NPs are known as “concealed question (CQ) NPs”. We first report the results of a pilot study that establishes co-occurrence patterns of the target embedding predicates. We then present two structural priming studies which test CQ NPs with “overt embedded questions”: “embedded wh-questions” and “embedded declaratives”. Both written sentence completion tasks demonstrate structural priming, which turns out to be sensitive to the phrase-clause distinction.

Key words: structural priming, concealed questions, embedded clauses, interrogative, declarative

Introduction

This paper investigates whether structural priming is sensitive to the phrase-clause distinction by testing priming for concealed questions (CQs).

Structural priming

Structural priming (Bock 1986) is the facilitating effect of experience with a certain morphosyntactic form on future processing. For example, when participants are asked to describe an event that could be expressed with either a prepositional object (PO) (e.g. ‘to a little boy’ in (1)) or a direct object (DO) (e.g., ‘a little boy’ in (2)), they produce more PO utterances after a PO prime (1) than after a DO prime (2), and more DO sentences after a DO prime.

(1) The man gave a new toy **to a little boy**.

(2) The man gave **a little boy** a new toy.

The same results hold for the active/passive alternation (Bock, 1986). Priming has also been demonstrated with different methodologies such as written sentence completion tasks (Pickering and Branigan 1998).

Concealed questions

Certain noun phrases (NPs) in English and other languages encode the meaning of a question without taking the form of one. These are aptly called concealed questions (CQs). In a nutshell, CQs involve shifting the meaning of a functional NP argument (3a) to a question-like meaning (3b) when that

NP appears as the object of certain embedding predicates (Heim 1979, Nathan 2006, a.o.).

(3a) *The committee announced [the winner of the award].*

(3b) *The committee announced [who had won the award].*

It has been argued that semantically CQs may not have the same denotation as a plain matrix interrogative (i.e., a set of propositions), but may instead behave more like an embedded declarative *that*-clause conveying the complete true answer to the question (i.e., a proposition) (Nathan 2006). Hence (4a) may be better paraphrased as a declarative (4c) rather than as an interrogative (4b).

(4a) *Kim told me [the capital of France].*

(4b) *Kim told me [what the capital of France was].*

(4c) *Kim told me [that Paris is the capital of France].*

Experiments

The goal of our study was twofold: to investigate possible priming with CQs and to use the experimental data to distinguish between the competing analytical approaches to CQs, as shown in (4b) vs. (4c). We report three experiments: a pilot study and two priming studies.

Pilot study

We conducted an offline cloze test with 24 sentence fragments to examine the natural frequencies of co-occurrence of CQ NPs and overt embedded questions (OECs).

(5) *The observer reported*

Experimental fragments contained two sets of embedding matrix verbs: *explain, learn, find out, figure out, report, disclose* (Set 1); *predict, estimate, determine, discover, announce, guess* (Set 2).

The pilot study (N=19) provided us with a profile for each individual verb. In terms of frequencies, most completions include *that*-clauses (38.33%), but there is also a significant number of NPs (31.67%). The percentage of embedded *wh*-questions is the lowest (11.11%).

Priming study 1: NPs and embedded questions

In the paper-based comprehension-to-production priming experiment, 44 participants were handed 4-page booklets with 60 items (12 Prime-Target pairs and 36 fillers). They read the prime sentences and completed the matching target fragments. Half of the primes contained CQ NPs (6a), and the other half contained embedded questions (EQs) (7a). Target fragments (6b) and (7b) were to be completed by the participants.

(6a) *The jeweler explained the value of the necklace.* (CQ NP-Prime)

(6b) *The engineer estimated* (Target)

(7a) *The expert explained what the diamond was worth.* (EQ-Prime)

(7b) *The technician estimated* (Target)

Participants produced more NP completions than EQ completions (significant main effect of **Completion Type**: $F=31,515$; $p=.00$; $\eta^2=.429$), consistent with the results of the pilot study.

There was also a significant two-way interaction between **Prime Type** and **Completion Type**: $F=5,398$; $p<.05$; $\eta^2=.114$. Participants produced more NP-responses after NP-primes than after EQ-primes, and more EQ-responses after EQ-primes. Therefore, despite the general high frequency of NPs, there is priming in the alternation between NPs and embedded-questions. Due to frequency, this priming is not symmetrical and seems to work better for NPs than for questions.

Priming study 2: NPs and embedded declaratives

On a theoretical plane, it is unclear whether CQs correspond to interrogative (4b) or declarative embeddings (4c). To investigate the distribution of clausal primes, the previous priming study was conducted using embedded declaratives (EDs) (8) instead of EQ primes (7a).

(8) *The expert explained that the diamond was worth five thousand dollars.*

In a web-based adaptation of the same task as in Priming Study 1¹, 20 participants read and completed the sentences.

There is a significant main effect of **Prime Type** ($F=6,211$; $p<.05$; $\eta^2=.257$). This means that whether the prime is nominal or clausal has an effect on the completions overall: CQ NPs prime NPs, and EDs prime clausal completions. In addition to the nominal and clausal completions, we also have a third type of completion, which is neither an NP nor an ED, with *that*:

(9) *He remembered ...how to write...*

There is a significant interaction between **Prime Type** and **Completion Type** ($F=5,492$; $p<.05$; $\eta^2=.234$) as in the previous study. This suggests that participants provided more NP completions after NP primes than after ED primes and more ED completions after ED primes than after NP primes. The priming effect is confirmed for the CQ NP-ED pair, as well.

Discussion

Concealed questions have an effect on structural priming when paired with overt embedded questions, expressed either as embedded interrogatives or embedded declaratives. These results provide novel evidence for structural priming being sensitive to the distinction between phrases and clauses.

Researchers using primary data have proposed that concealed questions could be interpreted as embedded wh-questions or as embedded polar interrogatives. In our study, a large proportion of the ED responses would be more appropriate as an answer to a yes-no-question than to a wh-question (e.g., *The police disclosed that the suspect was found.*). This suggests that while EDs prime EDs, their denotation may be more general than the one

that corresponds to the denotation of CQs. Thus, our results underscore the contrast between wh-questions and yes-no-questions and indirectly suggest that the wh-question interpretation may be more appropriate for CQs.

Notes

1. We also adapted CQ NP-EQ priming study to an internet-based format. As previously, there is a significant main effect of answer type ($F=15,725$; $p<.01$; $\eta^2=.466$) with more NP completions than EQ completions. However, here we don't get the significant prime type*completion type interaction. This could be due to the differences in methodology or to the smaller sample size (20 vs. 44).

Acknowledgements

We acknowledge TÜBİTAK and Harvard University FAS for their support. We are grateful to I. Caponigro, P. Graff, A. Hohenberger, E. Kravtchenko, A. M. Morgan, E. R. Su, and participants of the 6th International Workshop on Language Production in Edinburgh for their valuable help and comments.

References

- Bahadır, G., Polinsky M. 2010. Structural Priming of CQs. Poster presented in the 6th International Workshop on Language Production, Edinburgh, UK.
- Bock, J.K. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18, 355-387.
- Caponigro, I., Heller, D. 2007. The non-concealed nature of free relatives: Implications for connectivity in specificational sentences. In Barker, C., Jacobson, P. (eds.), *Direct Compositionality*, 237-263. Oxford, UK, OUP.
- Heim, I. 1979. Concealed Questions. In Bäuerle, R., Egli, U., von Stechow, A. (eds.), *Semantics from different point of views*, 51-60. Berlin, Springer-Verlag.
- Nathan, L. 2006. On the interpretation of CQs. PhD Thesis, MIT, MA.
- Pickering, M.J., Branigan, H.P. 1998. The representation of verbs: Evidence from syntactic priming in language production. *JML* 39, 633-651.

Syntactic recursion and theory-of-mind reasoning in agrammatic aphasia

Zoltán Bánréti, Éva Mészáros

Research Institute for Linguistics, Budapest, Hungary

Abstract

This study investigates how aphasic impairment impinges on recursivity of language and theory-of-mind inferences. Results of linguistic tests showed that in Broca's aphasia syntactic recursion is substituted for theory-of-mind inferences on the base that linguistic system and theory-of-mind type reasoning interact with one recursion modul in human mind.

Key words: syntactic recursion, theory of mind, aphasia

Introduction

We focussed on empirical investigations involving linguistic tests administered to subjects with agrammatic and Wernicke's aphasia. The test sessions involved 5 aphasics, 3 Broca's, 2 Wernicke' aphasics as well as 21 healthy control subjects. All aphasic participants had a left unilateral brain lesion. Aphasic subjects were assigned to aphasia types on the basis of CT and Western Aphasia Battery (WAB) tests c.f. Kertesz (1982).

Methodology

Photographs representing situations of everyday life were presented to subjects and questions were asked about them. We used 208 photographs (Stark, 1998) for each test, administered in three sessions. Within the same session, no picture was involved in more than a single question type. The types of questions involved were as follows:

Type 1: *What is X doing in the picture?* The question does not require that any of its own constituents should be involved in the structure of the answer. Type 2: *What does X hate / like / want /... every afternoon / in her office etc.?* The answer should be structurally linked to the question and involve: (i) a subordinate clause in direct object role, introduced by a recursive operation and signaled by a subordinating conjunction, or (ii) the verb of the question and its infinitival direct object, or (iii) a definite noun phrase in the accusative. Type 3: *What can be the most entertaining/unpleasant/urgent thing for X to do?* The answer should be structurally linked to the question and involve: (i) a subordinate clause in subject role, introduced by a recursive operation and signaled by a subordinating conjunction, or (ii) a bare infinitive subject, or (iii) a definite noun phrase in the nominative. Type 4: *What can X say / think / remind Y of / ask Y to do etc.?* The

structurally linked answer must be a clause embedded introduced by a recursive operation and signaled by a subordinating conjunction.

Type 1 questions did not restrict the structure of the answer in any way. Type 2 and Type 3 questions allowed for recursive and non-recursive answers alike. Finally, Type 4 questions could only be answered in a structurally linked way by using an embedded clause, introduced recursively. The most important results of the linguistic tests can be summarised as follows.

Results

Answers given by the five aphasic subjects and ten control subjects have been classified in terms of whether (i) they were structurally linked to the questions and were or were not grammatical; or (ii) they were not structurally linked to the questions and were or were not grammatical.

Table 1. The ratio of structurally linked answers to all answers is given within the brackets; the percentage of grammatical answers is given outside the brackets.

Question Subjects	Wernicke's aphasics 2	Broca's aphasics 3	Normal control 21
Type 1 question Number of answers:	(68.8) <u>50.8</u> 107	(56.3) <u>45.6</u> 257	(100.0) <u>100.0</u> 1117
Type 2 question Number of answers:	(46.0) <u>40.3</u> 111	(37.1) <u>31.6</u> 246	(100) <u>99.1</u> 1115
Type 3 question Number of answers:	(60.4) <u>34.3</u> 120	(30.8) <u>22.8</u> 240	(100) <u>98.5</u> 1139
Type 4 question Number of answers:	(66.7) <u>61.1</u> 76	(60.3) <u>45.6</u> 218	(100.0) <u>100.0</u> 982

According to Table 1 the number of structurally linked and grammatical answers decreased from Type 1 to Types 2 and Types 3. With respect to Type 4 questions (*What does X say / think / remind Y of / ask Y to do?*), requiring a recursively embedded clause as an answer, the performance of the subjects actually turned out to be better than with Type 1 questions (*What is X doing?*); or it was almost as good.

With respect to **Type 4** questions Wernicke's aphasics produced some conjunction-initial clauses and some clauses involving the subjunctive (i.e., the mood directly indicating subordination). Broca's aphasics gave few

answers beginning with a subordinating conjunction. The majority of structurally linked *and* grammatical answers produced by Broca’s aphasics, as well as the rest of the answers given by Wernicke’s aphasics, were statements that assumed the point of view of one of the characters seen in the picture, rather than being purely descriptive. The subjects answered the question as if they were in the “mental state” of the characters. These answers are referred to as “situational statements” with ‘*theory-of-mind*’ type reasoning. In them, the verb was inflected in the first, rather than the third, person singular. They directly represented the thought or statement of the characters they “cited”. Most of them did not involve a subordinating conjunction. An example for situational statement:

The picture: *A girl is showing her scar to a boy.*

Question: *Vajon mire gondol a fiú? ‘What may the boy be thinking of?’*

S.T.’s answer: *Mindjárt rosszul leszek! ‘I’m going to be sick’.*

Possible recursive construction: *Arra gondol, hogy mindjárt rosszul lesz. ‘He thinks he is going to be sick.’*

Table 2. The share of situational statements in answers to Type 4 questions:

Category Subjects	Wernicke’s aphasics (answers :76)	Broca’s aphasics (answers: 218)	Normal control (answers: 982)
Situational statement	(43.3) <u>43.3</u>	(74.0) <u>60.3</u>	(31.0) <u>31.0</u>
Sentence with subjunctive mood	(14.2) <u>14.2</u>	(10.0) <u>10.0</u>	--
Subordinating conjunction + situational statement	(12.5) <u>12.5</u>	(14.5) <u>8.7</u>	(24.0) <u>24.0</u>
Subordinating conjunction + descriptive clause	(30.0) <u>30.0</u>	(1.4) <u>1.4</u>	(45.0) <u>45.0</u>

Discussion

The majority of Broca’s aphasics’ answers to Type 4 questions were simple situational statements containing theory-of-mind type reasoning without subordinating conjunctions. The low percentage of subordinating conjunction in Broca’s aphasic’s answers shows that syntactic structural recursion is impaired.

In the sense of Takano - Arita (2010) theory-of-mind type reasoning is recursive. Agrammatic Broca’s aphasics may use recursive theory-of-mind inferences (and situative sentences carrying them) in their responses as a repair/compensatory strategy in order to avoid syntactic-structural recursion.

The content of situational statements showed that Broca's aphasic subjects correctly identified themselves with the mental states of the characters in the pictures, c.f. Siegel – Varley - Want (2006).

The use of simple situative statements could also be observed in the case Acknowledgements of control subjects, but only in 31% of their responses.

The share of situational statements jumped up in Broca's answers to Type 4 questions. A subset of linguistic devices indicating non-descriptive perspective was available for subjects, c.f.: theory-of -mind statements were based on a very simple syntax, verbs inflected with the first person singular feature. The complex linguistic subsystem of syntactic recursion was partially available or was not available at all for aphasics. Syntactic structural recursion was substituted for theory-of-mind recursion on the base that syntax and social cognition system interact with one common recursion modul.

Conlcusion

Findings support a model that posits a recursion module in the human mind that is shared by linguistic and theory-of-mind performance. In agrammatic aphasia syntactic representations are disconnected from the recursion modul but theory-of-mind type reasoning can access to the recursion modul.

Acknowledgements

This research has been supported by the National Scientific Research Fund (OTKA), project: NK 72461.

References

- Kertesz, A. 1982. The Western Aphasia Battery. New York, Grune & Stratton.
- Siegal, M., Varley, R., Want, S.C. 2006. Mind Over Grammar. Reasoning in Aphasia and Developmental Contexts. In Antonietti, A., Liverta-Sempio, O., Marchettio, A. (eds), Theory of mind and language in developmental contexts, New York, Springer, 107- 119.
- Takano, M., Arita, T. 2010. Asymmetry between Even and Odd Levels of Recursion in a Theory of Mind, in: Proceedings of ALife X , ed.: L.M. Rocha, L.S. Yaeger, M.A. Bedau, D. Floreano, R.L. Goldstone, A. Vespignani <http://www.citeulike.org/user/jasonn/article/7293338>

Recursion in language, theory-of-mind inference and arithmetic: aphasia and Alzheimer's disease

Zoltán Bánréti, Éva Mészáros, Ildikó Hoffmann, Zita Órley
Research Institute for Linguistics, Budapest, Hungary

Abstract

We found a variety of dissociations among the linguistic and non-linguistic operations tested. In the case of Hungarian speaking patients with a medium-serious AD, syntactic recursion is relatively unimpaired, as opposed to their limited ability to tackle theory-of-mind and arithmetical recursion. Conversely, we found limited syntactic recursion but normal theory-of-mind inferences and recursive arithmetical operations in Hungarian speaking agrammatic aphasics. We argue for a model that posits a module of recursive operations in the human mind that is shared by linguistic, theory-of-mind, and arithmetical performance.

Key words: recursion, mind, aphasia, Alzheimer disease

Introduction

Hauser, Chomsky & Fitch 2002 claim that the human faculty of recursion is revealed by syntactic-structural embeddings. Siegal and Varley (2006) argue, on the basis of experiments involving aphasic speakers, that theory of mind abilities may remain unimpaired even in cases of limited language faculties, that is, they are not grammar-dependent. Theory of mind' type reasoning can be recursive (cf. Takano -Arita 2010).

Aphasics

Bánréti (2010) conducted tests involving Hungarian speaking aphasic speakers. The results showed that, while in agrammatic (Broca's) aphasia structural recursion is limited, theory-of-mind recursion may be selectively retained. In the test situation, photographs representing everyday situations (Stark, 1998) were presented to subjects ('*What is it that X in the picture might be thinking of?*'; '*What do you think Y in the picture might be asking Z to do?*'). Broca's and Wernicke's aphasics gave answers whose grammatical versions were one of two types: (1) sentences involving structural recursion prompted by the structure of the question ('*X is thinking that he would be late*'), (2) so-called situative sentences (e.g., the aphasic person impersonates X and cites her thought or state of mind as in (*Christ, I'll be late!*)). Situative sentences are non-descriptive; rather, they represent the "cited" person's thought or intention directly, in first person singular. The subjects take over the point of view of one of the characters in the picture, identify themselves with that character. It is an important fact that in such answers it is not necessary to use a syntactic structure involving clausal embedding; a simple or even fragmented clause that expresses the theory-of-mind inference is just

enough. 73.4% of Broca's aphasics' answers were of this kind. Consequently, they were able to avoid using syntactic-structural recursion, an operation with respect to which their abilities were demonstrably limited. The use of simple situative sentences could also be observed in the case of control subjects, but only in 41% of their responses. Therefore, structural and theory-of-mind recursion represent two alternative strategies of which members of the control group were able to choose at will, whereas the aphasics were forced to choose the use of situative sentences.

Subjects with Alzheimer disease

The validity of the above observation can be supported by cases exhibiting the converse dissociation. The same tests were conducted with the participation of the Hungarian speaking subjects with Alzheimer disease (AD). The subjects were classified on the basis of Mini-mental state examination (Folstein, Folstein, & McHugh, 1975, ADAS-Cog (Rosen et al, 1984). We administered the above pictures and questions (*'What might X in the picture be thinking of?'*; *'What might Y in the picture be asking Z to do?'*) to four persons with slight and two with medium-serious AD.

In the responses of subjects with slight AD, there was no significant difference between them and normal controls in the proportions of replies involving syntactic-structural recursion vs. situative responses.

In the case of medium-serious AD the ratio of situative statements was significantly **lower** than in normal control subjects. Medium-serious AD subjects tend to avoid utterances in first person singular that assume the state of mind of another person. The share of sentences involving syntactic-structural recursion (embeddings involving *'that'* clauses) was not lower than in normal control responses. However, 37 % of these responses were semantically irrelevant statements referring to some marginal aspects of the picture. That is, while syntactic-structural recursion may remain unaffected in medium-serious AD, theory-of-mind inferences seem to be impaired to some extent. In order to support this with additional evidence, we administered false belief tests to our two medium-serious AD subjects (following Youmans & Bourgeois 2010). The secondary false belief test prove to be difficult for the medium-serious AD subjects they gave the wrong answers. Results of the secondary false belief test support the limitation of theory-of-mind inference abilities for both medium-serious AD patients.

In sum: in the slight and medium-serious stages of Alzheimer's disease, recursive clausal embedding abilities remain unaffected, but recursive theory-of-mind inferences become limited.. Unlimited syntactic-structural recursion and limited theory-of-mind inferences in Alzheimer's disease as opposed to limited syntactic-structural recursion and unimpaired theory-of-

mind inferences in Broca's aphasia: this is a pattern of double dissociation. This finding supports theories (e.g., Zimmerer-Varley 2010) that argue for the mutual independence of these two types of recursion in adults.

Arithmetic

With respect to recursion in arithmetical calculations, we found another case of double dissociation. Following Varley et al. (2005), we gave seven different tasks to agrammatic aphasics and persons with medium-serious AD. With the latter, we found limitation in the recursion of arithmetical operations. This limitation did not concern the four fundamental operations with one-digit numbers. Our AD subjects did have an idea of numbers, and they were able to transpose numbers from verbal to visual representation, but the idea of the infinity of numbers and any operations based on it were not accessible. They were unable to produce a recursive structure (e.g., generating inserting parentheses into calculation tasks and figuring out the result) and the idea of infinity of numbers was not accessible.

The four Hungarian speaking agrammatic aphasics, apart from minor memory problems, did not exhibit any limitation in arithmetical operations, they calculated correctly, they were able to produce potentially infinite sequences of numbers, applied recursive arithmetical operations correctly, inserted parentheses in various combinations, even double ones, and came up with the correct results. They did all that using numerical symbols and operation signs, in other words, they were not necessarily able to verbalise their otherwise correct operations.

Conclusion

The accessibility of recursive operations is limited in Alzheimer's disease for theory-of-mind and calculation but unlimited with respect to linguistic representations; whereas in agrammatic aphasia, linguistic representations may be disconnected from the recursion module, whereas theory-of-mind and calculation systems are able to access it (cf. Zimmerer – Varley 2010). These dissociations argue for a theoretical model that posits a module of recursive operations in the human mind that is shared by linguistic, theory-of-mind, and arithmetical performance. This common recursion module is accessible to a limited extent for the theory-of-mind and arithmetical subsystems while it is fully accessible for representations of linguistic constructions in Alzheimer's disease, whereas in agrammatic aphasia, the representations of linguistic constructions may be detached from the recursion module while theory-of-mind and arithmetical systems may access it at will.

Acknowledgements

This research has been supported by the National Scientific Research Fund (OTKA), project: NK 72461

References

- Bánréti, Z. 2010. Recursion in aphasia, *Clinical Linguistics & Phonetics*, 2010, Vol. 24, No. 11 , 906-914.
- Folstein, M.F., Folstein, S.E., McHugh, P.R. 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research*, 12/3: 189-98.
- Hauser, M.D., Chomsky, N., Fitch, T. W. 2002 .The Faculty of Language: What Is It, Who Has It, and How Does It Evolve? *Science*, 298,1569–1579.
- Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *American Journal of Psychiatry*. 1984. Nov;141(11):1356-64
- Siegal, M. Varley, R., Want, S.C. 2006. Mind Over Grammar. Reasoning in Aphasia and Developmental Contexts . In Antonietti, A., Liverta-Sempio, O., Marchettio, A. (eds): *Theory of mind and language in developmental contexts*, Springer, 107-119
- Takano, M., Arita, T. 2010. Asymmetry between Even and Odd Levels of Recursion in a Theory of Mind, in: *Proceedings of ALife X* , ed.: L.M. Rocha, L.S. Yaeger, M.A. Bedau, D. Floreano, R.L. Goldstone, A. Vespignani www.citeulike.org/user/jasonn/article/7293338
- Zimmerer, V., Varley, R. 2010. Recursion in severe agrammatism, in: van der Hulst H. (Ed.) *Recursion and Human Language*, 393-405. de Gruyter Mouton.
- Youmans, G., Bourgeoi, M. 2010. Theory of mind in individuals with Alzheimer-type dementia. *Aphasiology* 24:4, 515-534.

Linguistic and non-linguistic investigation of motion events

Ayşe Betül Toplu¹, Deniz Zeyrek²

¹Dept of Foreign Language Education, Middle East Technical University, Turkey

²Dept of Foreign Language Education & Department of Cognitive Sciences, Middle East Technical University, Turkey

Abstract

The verb-framed vs. satellite-framed language dichotomy (Talmy, 1985) is the most common framework used in the crosslinguistic investigation of motion event expressions. According to the classification, French is a verb-framed language, it integrates the path of motion into the main verb and uses a separate component to express the manner of motion. On the other hand, English, a satellite-framed language, gives the manner information in the main verb and expresses the path of motion with a separate component. The present study uses this dichotomy to see whether the motion event expressions patterns of these two typologically different languages (French and English) are also reflected in their motion event categorisations or not.

Key Words: motion events, typology, language and cognition

Theoretical background

The modern and systematic treatments of motion from a linguistic perspective start with Talmy (1985/2007)'s influential work. Although all languages have ways of describing a motion event, they express the same event differently. Talmy splits languages of the world into two categories, *verb-framed languages* (V-languages) and *satellite-framed languages* (S-languages). French is a V-language where path is typically encoded in the verb and manner is mostly expressed with an adverbial, as in (1). English is an S-language and instead of encoding the core component of a motion event (path) into the main verb, it uses a 'satellite' (e.g. a particle). S-languages have a tendency to encode manner into the main verb, as in (2).

(1) Le jeune homme est descendu les escaliers en courant.

Eng. The young man descended the stairs by running.

(2) The young man ran down the stairs.

There are a great number of studies investigating the motion event expressions in different languages (e.g. Özçalışkan and Slobin 1999&2000, Papafragou *et al.* 2006, Pourcel and Kopecka 2006, Choi-Jonin and Sarda 2007). On the other hand, there is another line of research questioning the effect of those crosslinguistic verbalization differences on the conceptual representation of motion events. This question is part of the renowned linguistic relativity hypothesis (Whorf, 1956). Scholars who take motion as the testing ground for this line of research use both verbal and non-verbal

experiments and analyze the results crosslinguistically (e.g. Gennari *et al.* 2002, Papafragou *et al.* 2002&2006, von Stutterheim and Nüse 2003).

Aim and scope

The first aim of the present study is to test the Talmyan dichotomy on two typologically distinct languages, i.e. English and French, by using a verbal production task. The second aim is to relate the results of the first task to the language and cognition debate by using a non-verbal categorization task. The main question here is whether the English and French speakers' categorization performances are influenced by the motion event verbalization patterns of their respective languages.

Methodology

20 monolingual speakers of English and 22 monolingual speakers of French, aged 20-35, took part in the study. The stimuli used in the experiments were 60 real-life video sequences exclusively shot for this purpose. There were 12 actions (crawl, dance, hop on one foot, hop on two feet, limp, march, patter, run, stagger, tiptoe, whirl and zigzag), each depicted with 5 different directions (into, out of, up, down and across).

The subjects first took the categorization task, during which they were to watch 30 motion events. The videos were organized in groups of three. In each group, there was a main video and two candidate videos. First, they watched the main video and then the candidate videos. One of the candidates was a same-manner alternate of the main video and the other was a same-path alternate. At the end, they were asked to choose which one of the candidates is more similar to the main video. Then the same subjects took the verbal production task, during which they watched and described 25 other motion events.

Results

Results of the production task showed that English speakers used manner sentences in 83% of their descriptions, whereas French speakers use path sentences in 96% of their utterances. The One-Way ANOVA revealed a significant difference between the two language groups ($F(1,40)=623.844$, $p<.01$). On the other hand, the categorization data showed that speakers of English and French do not differ at all in a non-verbal task, both having a significant tendency towards choosing the manner component as the criterion of similarity (72% and 64%, respectively). It was hypothesized that their choice of manner or path as the criterion for similarity would be influenced by the dominant verbalization patterns in their languages, which was clearly not the case in our data.

Discussion

The results of the verbal production task are perfectly consistent with the verb-framed language vs. satellite-framed language dichotomy, French speakers dominantly using path sentences and English speakers using manner sentences while describing motion events.

The results of the second task, on the other hand, may very well be interpreted as suggesting a clear evidence for the *universal approach*, which suggests that linguistic representation and conceptual representation are independent of each other, and which is presented as a counter-argument to the linguistic relativity hypothesis. It may also present evidence for the *underspecification / underdeterminacy view*, which claims that the cognitive representation (of motion events) is not bound by the linguistic labellings of one's native language (Papafragou and Selimis, 2010).

Acknowledgment

This study is carried out as part of a Research Project (N°109K281) funded by The Scientific and Technical Research Council of Turkey (TUBITAK).

References

- Choi-Jonin, I., Sarda, L. 2007. The expression of semantic components and the nature of ground entity in orientation motion verbs: A cross-linguistic account based on French and Korean. In Aurnague, M., Hickmann, M., Vieu, L. (Eds.), *The categorization of spatial entities in language and cognition*, 155-203. Amsterdam: John Benjamins (Human Cognitive Processing Series).
- Gennari, S., Sloman, S., Malt, B., Fitch, T. 2002. Motion events in language and cognition. *Cognition*, 83, 49-79.
- Özçalışkan, Ş., Slobin, D.I. 1999. Learning how to search for the frog: Expression of manner of motion in English, Spanish, and Turkish. In Greenhill, A., Littlefield, H., Tano, C. (Eds.), *Proceedings of the 23rd Annual Boston University Conference on Language Development*, 541-552. Somerville, MA: Cascadilla Press.
- Özçalışkan, Ş., Slobin, D.I. 2000. Climb up vs. ascend climbing: Lexicalization choices in expressing motion events with manner and path components. In Catherine-Howell, S., Fish, S.A., Lucas, T.K. (Eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development: Vol. 2*, 558-570. Somerville, MA; Cascadilla Press.
- Papafragou, A., Massey, C., Gleitman, L. 2002. Shake, rattle, 'n' roll: the representation of motion in language and cognition. *Cognition* 84, 189-219.
- Papafragou, A., Massey, C., Gleitman, L. 2006. When English proposes what Greek presupposes: The linguistic encoding of motion events. *Cognition*, 98, B75-B87.
- Papafragou, A., Selimis, S. 2010. Event categorisation and language: A cross-linguistic study of motion. *Language and Cognitive Processes* 25, 224-260.
- Pourcel, S., Kopecka, A. 2006. Motion events in French: Typological intricacies, Unpublished ms. University of Sussex and Max Planck Institute for Psycholinguistics.

- Talmy, L. (1985/2007). Lexical Typologies. In T. Shopen (Ed.), *Language Typology and Syntactic Description, Second Edition, Volume III: Grammatical Categories and the Lexicon* (pp. 66-168). Cambridge University Press.
- von Stutterheim, C., & Nüse, R. (2003). Processes of conceptualisation in language production. *Linguistics*, 41-5, (*Special Issue: Perspectives in language production*), 851-881.
- Whorf, B. L. (1956). *Language, thought and reality: Selected writings of Benjamin Lee Whorf*. Ed: Carroll J. B., MIT Press, Cambridge, MA.

Semantic priming at the sentence level: causal vs. purposive *because*

Joanna Blochowiak¹, Gözde Bahadır²

¹Department of Linguistics, University of Geneva, Switzerland

²Department of Cognitive Science, Middle East Technical University, Turkey

Abstract

In the present study we use the property of some linguistic items to express more than one meaning in order to investigate whether there is semantic priming at the sentence level. To test such a priming effect we use the connective *because* that can express a causal or purposive relation. If there is a priming effect, then participants will be expected to use causal *because* following causal primes and purposive *because* following purposive primes more frequently. The preliminary results do not reveal a clear priming effect but a slight trend toward priming, which we will continue to investigate increasing the number of participants.

Key words: semantic priming, pragmatic relationships, connective *because*, causal *because*, purposive *because*

Introduction

Numerous studies found priming effects at the phonological (Schriefers et al., 1990), syntactic (Bock, 1986) and lexical levels (Meyer and Schvaneveldt (1971). However, there is less evidence for priming at the semantic and pragmatic levels. Recent work has demonstrated the online expectation-driven processing of pragmatic relationships between sentences (Rohde&Horton CUNY2010).

The present study is the first one to address the question of semantic priming at the sentential level. The aim is to investigate whether there are priming effects related to the semantic properties of connectives, i.e. the relations they convey.

We propose to use the connective *because* whose semantics allows the expression of two kinds of relations: causal (1a) and purposive (2a).

(1) *Cause* condition

- a. P: Sam failed his exam because he didn't work at all.
- b. T: John turned off the TV because ...
- c. ... he didn't like the show [causal completion expected]

(2) *Purpose* condition

- a. P: Mary read the article carefully because she aims at getting an A
- b. T: Sally turned off the radio because ...
- c. ... she wanted to sleep [purposive completion expected]

The ongoing experiment is designed as a comprehension-to-production priming study where the participants are asked to first read a prime sentence, expressing *cause* in one condition (1a) and *purpose* in the other (2a), and then to complete a target fragment containing the connective *because* (1b/2b). Crucially, the target sentence can have both continuations: causal or purposive. If there is priming, we expect more causal completions (1c) following the cause-prime and more purposive completions (2c) following the purpose-prime.

Material and methods

14 adult (mean age: 27) participants (9 American native speakers and 5 advanced speakers) took part in the study which was an Internet-based written completion task. The participants were presented with a list of sentences containing 6 causal primes (C-primes), 6 purposive primes (P-primes) and 3 fillers between each prime-target pair. Some of the sentences (prime targets and other) were not finished and the participants were asked to invent completions for unfinished sentences.

So, as for now, we have 168 cases to analyze (12 sentences per 14 participants) divided into 2 sub-groups: 84 answers for causal-primes (C-primes) and 84 for purposive-primes (P-primes).

Results

The repeated ANOVA results indicate that there is a significant main effect of answer type: $F(1, 13) = 68.002, p = .000, \eta^2 = .840$, which means that participants provided more causal completions than purposive completions overall irrespective of the prime type they read. But the two-way interaction (prime type*answer type) is not significant which means that at this stage no clear priming effect was found. Nonetheless, it seems that there is a very slight trend towards priming as the figure 1 shows:

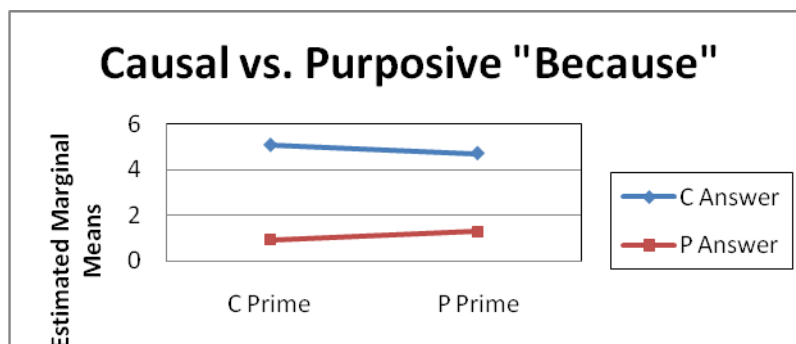


Figure 1. The plot showing the proportion of C- and P-answers to C- and P-primes respectively

Thus, since the experiment is still going on and the number of participants is not big enough, it is worth discussing these preliminary results on raw numbers in order to present our experimental assumptions transparently.

As the first step of our analysis we propose to identify how often in general people use *because* to express purposive relations with respect to causal ones. To this end we define a quantity of relative frequency of purposive *because* (FPB).

The FPB without priming can be obtained from corpus analysis. Our preliminary study of the corpus of journal “The Economist” indicates that people use the connective *because* to express purposive relations about 10 times less frequently than to express causal relations (for 100 occurrences of *because*, 8 cases of purposive uses have been found).

From our experiment we got two FPBs values: 15% (13 cases) with the C-priming and 21% (18 cases) with the P-priming. Thus the results of the experiment show that the average use of purposive relations FPB is increased from 15% up to 21% (18 cases) in presence of purposive primes (P-primes).

From the general rules governing priming we can make a conjecture that as the causal use of *because* is basically the norm and its purposive use is relatively rare, the C-primes have practically no effect and it is the P-primes which really affect the data. In this case we have a 6% increase of FPB (from 15% to 21%).

It is a promising outcome, suggesting manifestation of a priming effect even if the statistical results do not reach significance and therefore do not allow drawing a strong conclusion about the existence of these priming effects at the moment.

Another explanation would be that despite of rarity of purposive use of *because*, C-primes and P-primes could be equally strong. This hypothesis is less probable and will not be discussed here.

Moreover, our experiment and the proposed analysis points to another question: why the percentage of use of purposive relations differs between the result from the corpus study (10%) and the result we obtained from averaging the number of P-answers in presence of both types of primes (18%). The answer we would like to suggest is that this difference is the result of priming that persists over fillers. As we saw from the corpus, expressing purpose with *because* is not common and participants having become aware of this possibility through P-primes, start to exploit this option more frequently.

Conclusions and future developments of the experiment

The presented study is the first one addressing the issue of the semantic priming at the sentence level. Even if the preliminary results have not shown clear priming effects, on the basis of the data at hand we can observe a tendency suggesting the possibility of purposive priming. In order to verify if this tendency is confirmed we continue gathering data on the one hand and on the other, we consider taking some additional measures to test our hypothesis.

One of these measures concerns the FPB factor. For the moment we assume that the corpus' FPB provides a good enough approximation of the FPB without priming but ideally in order to measure FPB factor we should perform an experiment using our sentences in the absence of any priming context, which we are planning to do soon.

References

- Blochowiak, J., Bahadır, G. 2010. Semantic Priming at the Sentence Level: Causal vs. Purposive 'Because'. Poster presented in the 6th International Workshop on Language Production, Edinburgh, UK, 2-4 September, 2010.
- Bock, J.K. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18, 355-387.
- Meyer, D.E., Schvaneveldt, R. W. 1971. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, 90, 227-234.
- Rohde, H., Horton, W. 2010 Why or what next? Eye movements reveal expectations about discourse direction. Talk presented at the 23rd Annual CUNY Conference on Human Sentence Processing. New York, NY, March 18-20.
- Schriefers, H., Meyer, A. S., Levelt, W. J. M., 1990. Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language* 29, 86-102.

Subject gaps in German coordinative structures – Empirical evidence for a gradient phenomenon

Petra-Kristin Bonitz, Anke Holler

Department of German Studies, University of Goettingen, Germany

Abstract

In this article we investigate German complex clauses containing a subject gap using the empirical method of Magnitude Estimation. We will present evidence for the fact that subject gaps in coordinative structures can be characterized by (i) gradience and (ii) regional distinctiveness. Against the background of the recent critical discussion of Magnitude Estimation (e.g. Featherston 2008), we show that this method qualifies to test hierarchical graduation of acceptability.

Keywords: magnitude estimation, ellipsis, coordination, gradience, descriptive grammar of German

Phenomenon

The focus of our study lies on German complex coordinated clauses containing a subject gap as exemplified in (1).

- (1) *Wahrscheinlich geht Tino in die Kneipe und*
Likely goes Tino_{NOM} into a pub and
bestellt ein großes Bier.
orders a large beer.
'Tino probably goes into a pub and orders a large beer.'

It is usually assumed that this phenomenon, called SLF¹-coordination (Höhle 1983), has the following characteristics: (i) inversion at the first conjunct; (ii) subject gap (ellipsis) in the second conjunct; (iii) finite verb under C° by symmetric coordination (finite-frontal) or even asymmetric coordination (cf. Reich 2009). Thus, the linear structure of examples like (1) can be simplified in the following way (with 'e' marking a gap):

- (2) a. prefield V_{fin} NP_{Nom}[...] and e V_{fin} [...]
b. prefield V_{fin} NP_{Nom}[...] and e V_{fin} e [...]

Both the acceptability of coordinate constructions like (1) and their theoretical status are still under debate. The more normatively oriented Duden grammar (2009), for instance, only gives a tendential judgement about the preference of these structures (with or without the subject in the second conjunct). In contrast, different analysis of SLF-coordination, for example concerning the correct localization of the subject gap, have been proposed in the grammar-theoretical literature: Either the gap has been taken

to be located before the finite verb under SpecCP like in (2a) (cf. amongst others Fanselow 1991) or after the finite verb like in (2b) where the first gap marks the empty prefield (cf. amongst others Hartmann 1994). However, a comprehensive empirical assessment of SLF-coordination that makes predictions for an adequate modeling of the phenomenon is still outstanding.

The empirical data we present here allow a broad evaluation of the acceptability of SLF-coordination as well as of the proposed theoretical analysis to account for this phenomenon. We will show that SLF-coordination is a gradual and regionally variant grammatical phenomenon, which fact explains the wide spectrum of asserted judgements and theoretical accounts.

Method

Although there are many possible methods to collect data about linguistic structures such as corpus data, reaction time measurements, eye tracking, etc., we employed an opinion survey via Magnitude Estimation (ME) (Cowan 1997, McGee 2003) because of the advantages of this method for clarifying our specific problem. The supposed gradience and regional distinctiveness of SLF-coordination can be well detected with ME since this method is able to provide a hierarchical graduation of the acceptability of clause structures (see among others Bard et al. 1996, Schütze 1996). The merit of ME lies in the fact that one does not need a pre-defined scale, like a 5- or 7-point scale, which is limited by its very nature. With this method the test persons create their own magnitude scale by estimating sentences relative to each other so that even small differences in the acceptability of structures are documented.

The study has been realized as an online survey with 623 probands that had to judge the acceptability of stimuli of the following kind:

- (3) *Vielleicht fuhr [die Frau] in die Stadt und*
 Maybe drives [the woman_{NOM}] into the town and
[die Frau] besuchte Freunde.
 [the woman_{NOM}] visits friends.

‘Perhaps the woman drove into town and the woman visited some friends.’

We developed a systematic pattern that consists of eight scenarios which all share the same sentence types. These types were established in the way that the NP_{NOM} in the first and/or the second conjunct was manipulated so that 13 different conditions in each scenario emerged. The test subjects were German native speakers from Germany, Austria and Switzerland.

Results and discussion

By means of the collected data we argue that there is a linear distribution in the acceptability of complex clauses with realized subjects in both conjuncts versus complex clauses with a subject ellipsis in the second conjunct.

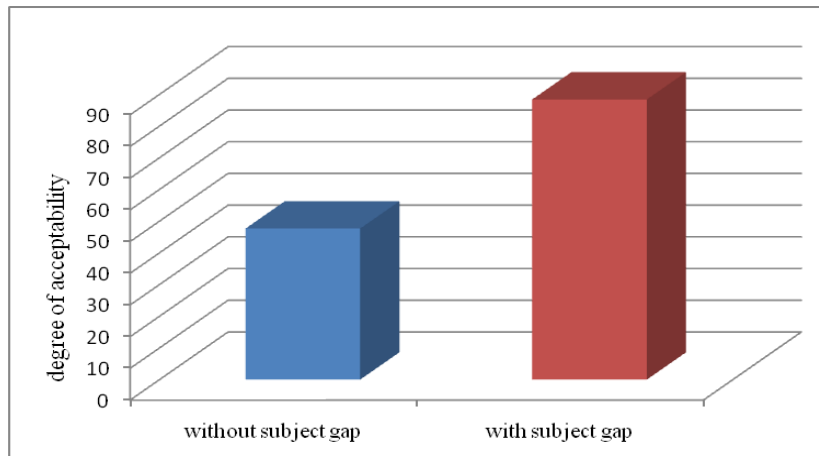


Figure 1. Distribution of the preferences.

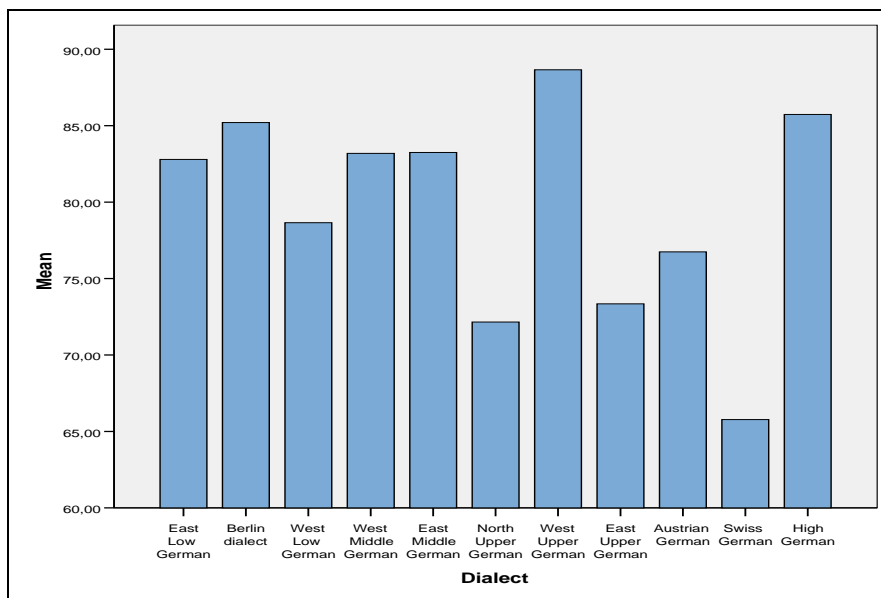


Figure 2. Regional effects.

Our results prove that subject ellipses are significantly preferred, which was shown by a T-test ($t(622) = 17.17, p < .05$) (cf. figure 1). This is contradictory to the assumptions in the normative German grammar (e.g. Duden 2009). By analyzing the data more specifically effects become obvious that give evidence for (i) gradual dependencies and (ii) differences throughout the regional German dialects. There are several conditions, such as the substitution of the noun phrase with a personal pronoun or an indefinite pronoun, that altogether support the gradual character of SLF-coordination. Its regional distinctiveness follows from observed regional differences. As is visualized in figure 2, test persons from the northern German speaking regions judge the elliptical construction better than test persons from the southern German speaking regions (especially the Swiss German). Because of the gradual and additionally regional dependent character of the SLF-coordination existing grammatical theoretical approaches have to be revised.

Summary

Crucial for the descriptive grammar is our result that empirical data show differences in the acceptability of German SLF-coordination. These structures can be characterized by gradience and regional distinctiveness.

Notes

1. *Subjektücke in finit-frontalen Sätzen* (subject gap in finite-frontal sentences)

References

- Bard, E.G., Robertson, D., Sorace, A. 1996. Magnitude Estimation of Linguistic Acceptability. *Language* 72, No.1, 32-68.
- Cowart, W. 1997. *Experimental Syntax. Applying Objective Methods to Sentence Judgments*. Thousand Oaks etc.: Sage.
- Dudenredaktion (ed.) 2009. *Duden. Die Grammatik*.⁸ Mannheim u.a., Dudenverlag.
- Fanselow, G. 1990. Minimale Syntax. *GAGL* 32.
- Featherston, S. 2008. Thermometer judgements as linguistic evidence. In Riehl, C. and Rothe, A. (eds.) *Was ist linguistische Evidenz?*, 69-89. Aachen, Shaker.
- Hartmann, K. 1994. Zur Koordination von V-2 Sätzen. *Zeitschrift für Sprachwissenschaft* 13, 3-19.
- Höhle, T. N. 1983. Subjektücken in Koordinationen.
(www.linguistik.uni-tuebingen.de/hoehle/manuskripte/SLF-W5.1_neu.pdf).
- McGee, M. 2003. Usability Magnitude Estimation. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*, 691-695.
- Reich, I. 2009. *Asymmetrische Koordination im Deutschen*. Tübingen, Stauffenburg.
- Schütze, C. T. 1996. *The Empirical Base of Linguistics. Grammaticality Judgments and Linguistic Methodology*. Chicago/ London, University of Chicago Press.

Prosody and quantifier semantics in Greek

Antonis Botinis¹, Aikaterini Bakakou-Orphanou², Anthi Chaida¹

¹Lab of Phonetics & Computational Linguistics, University of Athens, Greece

²Department of Linguistics, University of Athens, Greece

Abstract

This study investigates prosody and semantic specification in Greek with reference to affirmative and negative sentence quantifier functions. A production and a perception experiment were carried out, the results of which indicate: (1) The negative interpretation of the word ['liya] (few) is produced with focus on the quantifier, whereas the positive one with focus on the preceding verb. (2) The word in focus has an enlarged tonal range associated with the stressed syllable, while the material out of focus is tonally compressed. (3) The stressed syllable of the word in focus, especially the vowel, is significantly longer. (4) The affirmative and negative functions of the quantifier had high identification rates, indicating a one-to-one production and perception relation.

Key words: prosody, intonation, quantifiers, Greek.

Introduction

This paper reports on an experimental study of prosodic production and perception of semantic specification with reference to sentence quantifiers in Greek. The word ['liyo] (Sing.) and ['liya] (Pl.) may have a positive or a negative meaning, which is determined by prosodic structure. E.g. the sentence ['exume 'liya 'xrimata] may mean either “we have some money” or “we have little money”. We assume that a sentence with this quantifier may have different interpretations, which are determined by different focus applications. Focus is mostly associated with semantic weighting and information structure. The prosodic structure of focus with regards to the distinction new-old or the most important information unit in a context has been widely studied in many languages (e.g. Botinis 1989, Bruce 1977, Gussenhoven 2008), while in a recent study (Botinis et al. 2011), focus has been reported for the communicative distinction of agreement-disagreement.

Experimental methodology

In accordance with one *production experiment*, the prosodic structure of the word ['liya] with distinct semantic specification was examined, i.e. positive or negative, in the context of the carrier sentence [mu 'minane _ le'moja] (I have a few lemons left). Ten female speakers, in their early twenties, with standard Athenian pronunciation, produced the above speech material 2 times (normal tempo). The experimental sentences were answers of the leading question [θa mu 'fcaksis mpa lemo'naða] “Will you make me a lemonade?”. The speakers were instructed to answer either positively or

negatively. The recordings took place at Athens University Laboratory of Phonetics and Computational Linguistics. The speech material was analysed with Praat, and measurements were taken for each segment duration of the word ['liya] as well as 2 F0 points per syllable for each utterance.

In accordance with one *perception experiment*, 10 productions for each interpretation (negative/positive) were presented in random order to 20 Greek native speakers (10 males, 10 females), who were instructed to categorise each stimulus as semantically negative or positive.

Results

The results of the *production experiment* are shown in Figures 1-3.

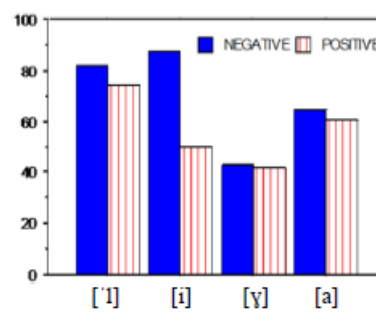
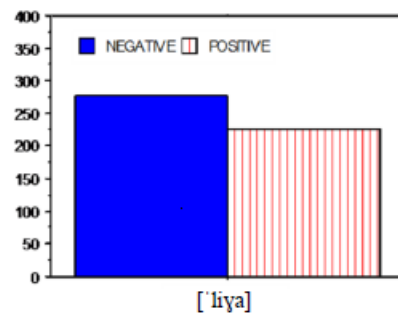


Figure 1. Word duration (ms) of the quantifier with negative (focus) and the positive (no focus) interpretation. Figure 2. Segmental durations (ms) of the quantifier with negative (focus) and positive (no focus) interpretation.

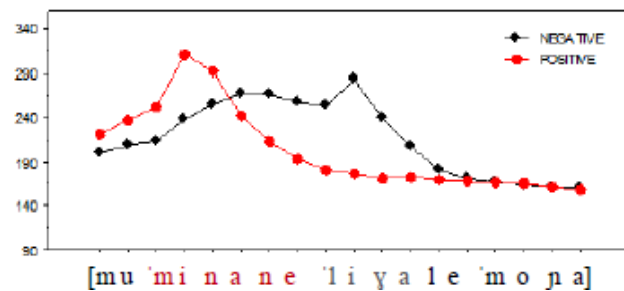


Figure 3. Mean F0 measurements (Hz) with negative semantic interpretation (quantifier ['liya] in focus) and positive one (verb ['miname] in focus).

The average word duration (Fig.1) of the quantifier when in focus (negative interpretation) is 277 ms, and out of focus (positive interpretation) 226 ms. Focus has a lengthening effect of 50 ms (18,3%); this difference though does not reach the significance level of 0.05 ($t(df\ 30)=0,45$, $p=0,65$). The consonant [l] in focus (Fig.2) is 81 ms and out of focus 74 ms (7,7 ms, 9,4% difference). This difference is not statistically significant at 0.05 level

($p=0.15$). The duration of the vowel [i] in focus is 87 ms and out of focus 50 ms (37,1 ms, 42% difference); this difference is highly significant ($p<0,0001$). Neither the consonant [ɣ] nor the vowel [a] show any worth mentioning difference in focus and out of focus and, thus, focus application has no effect on unstressed syllables. The mean duration for the first syllable ['li] when in focus is 169 ms, and out of focus 124 ms (44,8 ms, 26,5% difference); this difference between the two interpretations for the first syllable is highly significant ($t(df\ 30)=3,90$, $p<0,0005$). Neither the unstressed syllable [ya] nor the syllabic constituency segment show any worth mentioning difference.

With regards to intonation (Fig.3), 2 very distinct tonal structures are observed: (a) with focus on the quantifier for the negative interpretation, and (b) with focus on the verb for the positive interpretation. In rendition (a), a prominent rise-fall tonal movement is produced, with its peak aligned with the vowel of the stressed syllable of the quantifier ['liya] in focus. In rendition (b), a prominent rise-fall movement is also produced, with its peak aligned with the vowel of the stressed syllable of the verb ['minane] in focus. In both cases, the post-focal part of the utterance is tonally compressed, highlighting the respective word in focus. The tonal distinction is statistically significant ($F(1,574)=9,28$, $p=0.002$).

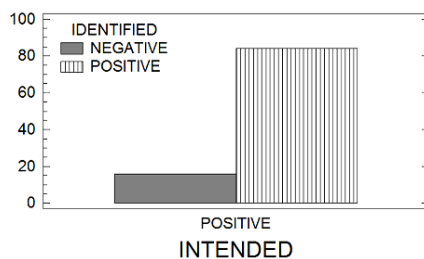


Figure 4. Identification (%) for intended positive interpretations.

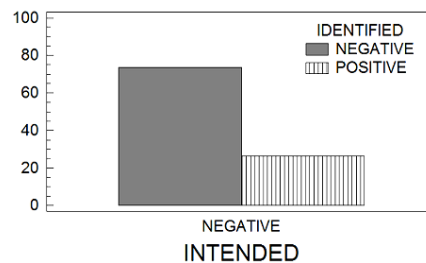


Figure 5. Identification (%) for intended negative interpretations.

The results of the *perception experiment* are shown in Figures 4-5. The positive interpretation (Fig.4) was identified as intended 84% of times, and the negative interpretation (Fig.5) 73% of times. The identification results in total were statistically significant ($\chi^2(df\ 1)=131,41$, $p<0.01$, 99% conf. level).

Discussion

Focus is a complex linguistic concept, which has been studied from many different perspectives. With reference to phonetic studies, several local and global correlates have been reported, such as local tonal expansion, local tonal alignment and global compression, especially a postfocal tonal

flattening (e.g. Botinis 1989, Hirst, Di Cristo 1998, Xu et al. 2005). According to the results of this study, the typical tonal structure of a tonal expansion aligned with the stressed syllable of the word in focus, in combination with a global tonal compression, especially a post-focal one, has been corroborated. However, the segmental material of the stressed syllable of the word in focus, especially the vowel, has a substantial lengthening effect. Given that this effect has not been constantly reported in previous studies for Greek, we may assume that duration structure is a variable parameter of focus, whereas tonal structure is a constant one. In fact, in the present study, duration was found to be a significant prosodic correlate of focus with a quantification distinction.

In conclusion, this study highlights that focus prosodic structure has a distinctive function for quantification with regards to sentence affirmation and negation in Greek. These distinctions were clearly perceived by listeners, indicating a one-to-one production and perception relation.

Therefore, we might assume a one-to-many relation between focus prosodic structure and semantic interpretation. According to this hypothesis, focus production may have only one prosodic structure, which may be related to several interpretations. Apparently, information focus reported in several studies of Greek prosody (e.g. Botinis 1989) and semantic specification of quantification reported in this paper do share the basic prosodic characteristics, at least tonal ones. However, focus application on the quantifier under examination has a sentence negation function rather than a highlighting one, which puts new dimensions to current prosodic knowledge and linguistic theory in general.

Acknowledgements

This research is supported by the University of Athens “Kapodistrias” project. Thanks to the participants of the experiments, to Marios Fourakis for his feedback, and to Sonia Loui for her help in the perception experiment.

References

- Botinis, A. 1989. *Stress and Prosodic Structure in Greek*. Lund University Press.
- Botinis, A. Bakakou-Orphanou, A., Chaida, A. 2011 (to appear). Sentence functions and tonal semantics in Greek. 7th AISV National Conference, Lecce, Italy.
- Bruce, G. 1977. *Swedish Word Accents in Sentence Perspective*. Lund, Gleruup.
- Hirst, D., Di Cristo, A. (eds). 1998. *Intonation systems: a survey of twenty languages*. Cambridge University Press.
- Gussenhoven, C. 2008. Notions and subnotions in information structure. *Acta Linguistica Hungarica* 55, 381-395.
- Xu, Y., Xu, C.X. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33, 159-197.

Phonology and phonetics of Greek palatalisation

Antonis Botinis¹, Anthi Chaida¹, Evgenia Magoula²

¹Lab of Phonetics & Computational Linguistics, University of Athens, Greece

²Department of Education, University of Athens, Greece

Abstract

The present investigation examines palatal production as well as the relation of phonology and phonetics in Greek. In accordance with one production experiment, the results indicate: (1) palatal productions that surface from one underlying segment are significantly longer than palatal productions that surface from two underlying segments, (2) palatal productions are voice assimilated with the preceding stop and fricative consonant production and, (3), the locus frequencies of male palatal productions are in the area of 1800-2000 Hz.

Keywords: acoustics, phonology, phonetics, palatalisation, Greek.

Introduction

The present paper reports on an investigation of palatalisation production in Greek in regard to phonology-phonetics interface and the relation between underlying phonological representations and surface phonetic ones. For example, the plural of neutral nouns is formed by the addition of the monomorphemic vowel /a/ and, given that a great deal of neutral nouns end in the vowel /i/, production of palatalisation may be triggered in a considerable part of the Greek lexicon. Thus, words like /ku.'pi/ “oar” and /ku.bi/ “button” in singular turn into [ku.'pça] and [ku.'bja] in plural, respectively, as a result of the application of the disyllabification and palatalisation rules (see further examples in Table 1).

Palatalisation is a major issue and has widely been discussed with reference to the phonem inventory and the phonological structure of Greek (e.g. Newton 1961, Koutsoudas, Koutsoudas 1962, Householder 1964). Its acoustic characteristics have however hardly been investigated and our knowledge is thus based on intuition and the knowledge of language.

In this investigation, we have concentrated on two main questions with reference to (1) the duration of the palatal productions and (2) the quality of palatal productions in terms of voicing. Thus, basic acoustic data are obtained, bring to bear crucial empirical evidence at the relation between underlying and surface representations.

Experimental methodology

The speech material of the present investigation consists of a set of key words in the carrier sentence ['ipe ____ so'sta] (s/he said ____ correct). The key words consist of neutral nouns in singular vs. plural forms, where the latter are in palatalisation context (Table 1).

Table 1. Underlying phonological representations (left) and surface phonetic representations (right) of neutral nouns in non-palatalisation contexts in singular and palatalisation contexts in plural.

Phonological representation			Phonetic representation	
Singular		Plural	Singular	Plural
/ku'pi/	“oar”	/kupi #'a/	[ku'pi]	[ku'pça]
/ku'bi/	“button”	/kubi #'a/	[ku'bi]	[ku'bja]
/sku'fi/	“hat”	/skufi #'a/	[sku'fi]	[sku'fça]
/klu'vi/	“cage”	/kluvi #'a/	[klu'vi]	[klu'vja]
/ku'ki/	“bean”	/kuki #'a/	[ku'ci]	[ku'ca]
/pu'gi/	“purse”	/puçi #'a/	[pu'ji]	[pu'ja]
/pa'xi/	“fat”	/paxi #'a/	[pa'çi]	[pa'ça]
/tra'yi/	“goat”	/trayi #'a/	[tra'ji]	[tra'ja]
/pu'li/	“bird”	/puli #'a/	[pu'li]	[pu'la]
/pa'ni/	“clouth”	/pani #'a/	[pa'ni]	[pa'na]

Five female and five male speakers of standard Athenian pronunciation, at their twenties, were asked to produce the speech material at their natural normal tempo. The speech material was produced at the recording studio of Phonetics and Computational Linguistics Laboratory at Athens University and was analysed with the Praat software programme.

Duration measurements of consonants as well as measurements of formant transitions between consonants and vowel of the second syllable were taken. Furthermore, consonants in palatalisation contexts were classified as voiced or voiceless, in accordance with visual inspection of the acoustic signal. The results of duration and spectral measurements were subjected to statistical analysis with the StatView software package.

Results

The presentation of the results will be confined to a description of the main effects of palatal productions. A qualitative analysis is presented (Figure 1) along with a quantitative analysis of palatal production durations (Figure 2). In general, the application of the palatalisation rule and hence palatalisation production has intraconsonantal as well as syllabic context effects.

The intraconsonantal effects are in the first place related to the production of stop and fricative palatal consonants, which are voice-assimilated with the preceding consonant, according to which a concentration of energy in low frequencies, i.e. a typical voice bar, is evident in voiced assimilation whereas lack of low energy is basically evident in voiceless assimilation (e.g. Figure 1.1 and 1.2). The intraconsonantal effects are also related to segmental durations, according to which palatal productions that surface from two underlying segments are significantly longer (anova, $p < 0.0001$) than palatal productions that surface from one underlying segment (see Figure 2).

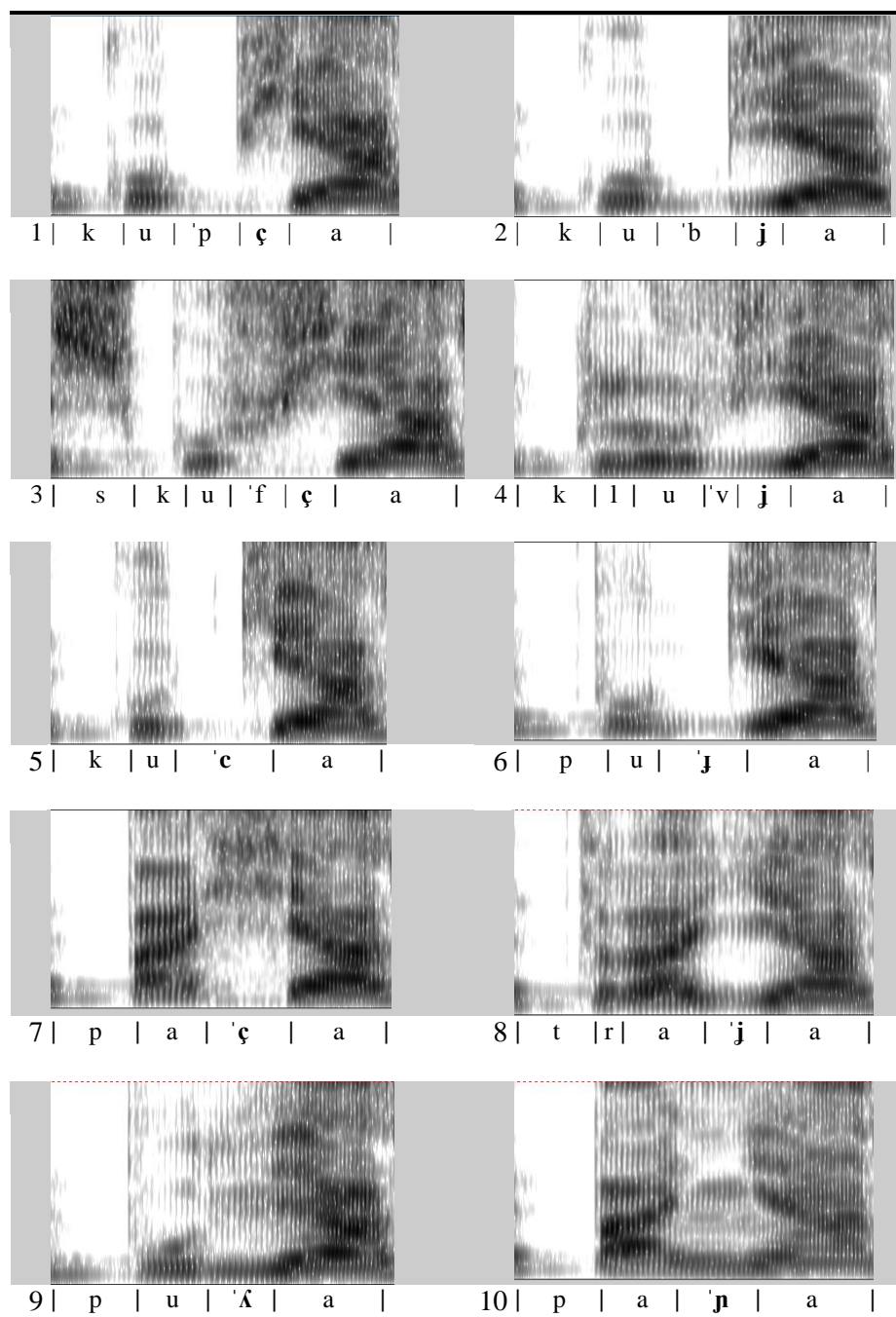


Figure 1. Exemplification of palatalisation productions (see text).

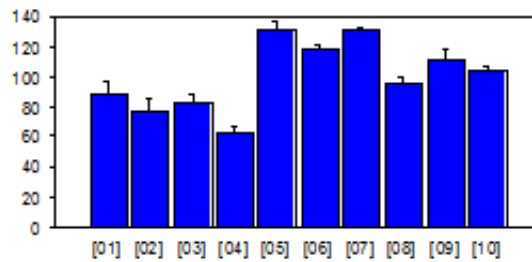


Figure 2. Segmental durations (in ms) of palatal productions (column numbers correspond to numbers of Figure 1).

The syllabic context effects are mainly related to formant transitions and locus frequencies with reference to the following vowel in the first place. Thus, in accordance with Figure 1, the locus frequencies of palatal productions are in the area of 1800-2000 Hz (male voice) whereas F1 and F2 have opposite slopes, i.e. a falling and a rising, respectively. Palatal productions between vowels have however a mirror-image acoustic structure (e.g. Figure 1.7, 1.8), which indicates that palatal production is related to intersyllabic acoustic structure, in addition to intraconsonantal one.

Discussion and conclusions

The results presented in the above section indicate that the production of palatalisation maps an underlying phonological structure into a surface phonetic one, in accordance with both temporal and spectral changes in the acoustic signal. Most importantly, palatal productions that surface from two underlying segments, i.e. a consonant and a vowel, are significantly longer than palatal productions that surface from one underlying segment, i.e. a vowel. A question arises whether this phonologically-triggered lengthening effect may override any duration effect of syllabic structure and the number of consonants in syllabic onset (Botinis et al. 1999), a hypothesis which remains to be tested. In general, in accordance with the results of the present investigation, the underlying phonological structure leaves acoustic traces in surface phonetic production and the acoustic signal, which may be related to both intraconsonantal as well as intersyllabic production, formant transitions from and to preceding and following vowels, respectively.

References

- Botinis, A., Fourakis, M., Prinou, I. 1999. Prosodic effects on segmental durations in Greek. Proceedings of the 6th European Conference on Speech Communication and Technology, Eurospeech 99, vol. 6, 2475-78, Budapest, Hungary.
- Newton, B. 1961. Rephonemization of Modern Greek. *Lingua* 10, 275-284.
- Koutsoudas, A., Koutsoudas, O. 1962. A contrastive analysis of segmental phonemes of Greek and English. *Language Learning* 12, 211-230.
- Householder, F. Three dreams of Modern Greek phonology. 1964. *Word* 20,3, 17-27 (Papers in memory of George C. Papageotes).

Perception of French, Belgian and Swiss accents by French and Belgian listeners

Philippe Boula de Mareüil¹, Alice Bardiaux²

¹ LIMSI – CNRS, Orsay, France

² Université Catholique de Louvain – FNRS, Louvain-la-Neuve, Belgium

Abstract

This article addresses the perceptual identification of French regional accents by listeners from the Paris region and Belgium. It is based on the geographical localisation of about thirty speakers from seven French-speaking areas: Vendée (West of France), Languedoc (South of France), Alsace (East of France), Romand Switzerland, East, Centre and West of Belgium. Contrary to the speakers' age bracket (older or younger than 60) and speaking style (read or spontaneous speech), listeners' region of origin and speakers' degree of accentedness (also rated by the listeners) have a major effect. Confusions are frequent among the Belgian areas, but taken as a whole, the Belgian accent is remarkably well identified, especially by Belgian listeners. The Southern accent remains identified best.

Key words: perceptual dialectology, regional variation, French accents.

Introduction

The way of speaking conveys a series of information from which attitudes and representations are built in speakers' mind. It is subject to cross- and within-speaker variation, as exemplified by regional accents. However, recent work based on perceptual experiments showed that fine-grained identification of regional accents is a difficult task, whether in American English (Clopper and Pisoni 2004) or in French (Woehrling and Boula de Mareüil 2006; Woehrling 2009). In particular, it was shown that confusions are frequent within Western French accents and within Southern French accents, irrespective of listeners' Western/Southern place of residence.

The present article expands upon previous experiments to investigate how various factors may impact the perceptual rating and/or identification of regional accents from France, Belgium and Switzerland. Some of these factors, assumed to be relevant in accent perception (Labov 1972; Preston 1989) are listeners' geographic origin, speakers' age and speaking style.

A perceptual test was conducted, involving listeners from France and Belgium. Based on speech samples from three areas of France, three areas of Belgium and one area of Switzerland, the test included a rating task and an identification task. The experimental corpus, subjects and tasks are described in the following section. Results are then presented and briefly discussed.

Experimental setup

The experimental corpus is made up of speech material recorded in Treize-Vents (Vendée, West of France, assumed to represent “standard” French), Douzens (Languedoc, South of France), Boersch (Alsace, East of France), Nyon (Canton de Vaud, Romand Switzerland), Liège (East of Belgium), Gembloux (Centre of Belgium) and Tournai (West of Belgium). The data were collected within the framework of the “Phonology of Contemporary French” (PFC) project (Durand *et al.* 2002). In each investigation point, four speakers were selected: one male and one female aged between 30 and 60, one male and one female aged over 60. For each speaker, a read sentence and an excerpt of spontaneous speech (of about 10 seconds each) were selected.

The 56 resulting stimuli were administered to 25 listeners from the Paris region and 25 listeners from French-speaking Belgium. The rating task consisted in evaluating the degree of accentedness on a 0–5 scale (0 = no accent; 5 = very strong accent) and the identification task consisted in a 7-alternative forced-choice categorisation. The experiment was conducted through a web-based interface (available at http://www.audiosurf.org/test_perceptif_cecile/).

Results

Average results for both listener groups are reported in Table 1. Degrees of accentedness are tabulated between square brackets. Identification scores are tabulated within the core of the confusion matrices.

On average, the accents were judged as rather strong (3.0/5 for French listeners and 2.7/5 for Belgian listeners). Even though the accents from Belgium were rated stronger by French listeners than by Belgian listeners, the overall difference is not significant between the degrees of accentedness assigned by the two groups of subjects, according to a *t*-test.

The average correct identification rate is 36% for French listeners and 44% for Belgian listeners. Analyses of variance (ANOVAs) were performed on listeners’ responses counted as right (1) or wrong (0) with the random factor Subject, the between-subject factor Group (French/Belgian) and two within-subject factors: speakers’ Age (± 60 years) and speaking Style (read/spontaneous). There is no significant effect of speakers’ Age and speaking Style, but there is here a significant effect of listeners’ Group [$F(1, 48) = 9.59$; $p < 0.01$]. A Degree of accentedness was also attributed to the stimuli by rounding the listeners’ average ratings, but no stimulus received an average rating close to 0. ANOVAs were thus conducted with a 5-level within-subject factor Degree of accentedness. This factor produced a major effect [$F(4, 196) = 51.9$; $p < 0.001$]. Results are similar when ANOVAs are conducted separately for each listener group.

Table 1. Degrees of accentedness [out of 5, between square brackets] and confusion matrices resulting from the answers of 25 listeners from (a) the Paris region and (b) Belgium. Percentages are given with respect to 200 answers. For each investigation point in Ve(nde), La(nguedoc), Al(sace), Sw(itzerland) and Be(lgium), majority answers are highlighted in boldface.

(a)

Results of French listeners							
Answer Origin	Treize- Vents	Douzens	Boersch	Nyon	Liège	Gembloux	Tournai
Treize-V. (Ve) [1.9]	38	5	13	10	16	11	8
Douzens (La) [4.8]	6	90	1	2	1	0	1
Boersch (Al) [3.7]	9	2	34	25	9	14	8
Nyon (Sw) [3.2]	4	2	20	38	14	12	11
Liège (Be) [3.0]	11	3	18	14	19	22	14
Gembloux (Be) [2.7]	18	4	14	17	16	19	13
Tournai (Be) [2.0]	28	3	14	10	13	20	13

(b)

Results of Belgian listeners							
Answer Origin	Treize- Vents	Douzens	Boersch	Nyon	Liège	Gembloux	Tournai
Treize-V. (Ve) [2.1]	37	12	10	6	10	19	6
Douzens (La) [4.2]	15	73	8	2	1	0	1
Boersch (Al) [2.8]	10	1	44	22	6	12	5
Nyon (Sw) [3.1]	14	5	18	55	2	2	4
Liège (Be) [2.7]	8	1	4	4	48	17	18
Gembloux (Be) [2.4]	4	3	5	2	22	39	25
Tournai (Be) [1.4]	23	6	4	2	9	36	20

Detailed results (given in Table 1) show that Douzens (Languedoc, here the only representative of the South of France) and Nyon (Switzerland) are identified best. The Liège, Gembloux and Tournai (Belgium) investigation points are often confused: their correct identification rate is 17% for French listeners and 36% for Belgian listeners. Yet, counting any of these three answers as correct, the identification rate as Belgian is 49% for French listeners and 77% for Belgian listeners. Even though responses are not very precise, the Belgian accent is thus remarkably well recognised.

Discussion

This study suggests that five French accents can be distinguished satisfactorily in France and its periphery: Western (standard) French, Southern French, Alsatian, Swiss and Belgian. Interestingly, identification was not affected significantly by speakers' age and speaking style. Degree of accentedness, however, was an influential factor: stronger accents were better identified than milder accents. Also, listeners' geographic origin was an influential factor, whereas it was not in Woehrling and Boula de Mareüil's (2006) experiment, based on listeners from Paris and Marseilles. Work is in progress to relate the results achieved to the phonetic characteristics of the accents under investigation.

Acknowledgements

We are grateful to the leaders of the PFC project Jacques Durand, Bernard Laks and Chantal Lyche. Our thanks also go to Cécile Woehrling, who designed and analysed the experiment conducted in the Paris region.

References

- Clopper C.G., Pisoni D.B. 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32, 111–140.
- Durand, J., Laks, B., Lyche, C. 2002. La phonologie du français contemporain : usages, variétés et structure. In Pusch, C. D. and Raible, W. (eds.), *Romance corpus linguistics — Corpora and spoken language*, 93–106. Tübingen, Narr.
- Labov, W. 1972. *Sociolinguistic patterns*, Philadelphia, University of Pennsylvania Press.
- Preston, D.R. 1989). *Perceptual dialectology*, Dordrecht, Foris Publications.
- Woehrling, C., Boula de Mareüil, Ph. 2006. Identification d'accents régionaux en français : perception et analyse, *Revue PArôle* 37, 25–65.
- Woehrling, C. 2009. *Accents régionaux en français : perception, analyse et modélisation à partir de grands corpus*, PhD thesis, Université Paris-Sud.

Evaluating speech samples designed for the Voice Profile Analysis Scheme for Brazilian Portuguese (BP-VPAS)

Zuleica Antonia de Camargo¹, Sandra Madureira¹, Luiz Carlos Rusilo²

¹Integrated Acoustic Analysis and Cognition Laboratory-LIAAC-Pontifical Catholic University of São Paulo, São Paulo, Brazil

² Department of Actuarial and Quantitative Methods-DAMQ-Pontifical Catholic University of São Paulo, São Paulo, Brazil

Abstract

The present study aims at evaluating a corpus designed for voice quality settings analysis (Laver, 1980). It addresses the role of keysepeech segments for the identification of voice quality settings. The corpus, recorded by 60 speakers, contained repetitions of 10 keysepeech sentences, and semi-spontaneous speech samples. These samples were evaluated by two expert subjects according to the BP-VPAS Profile (Camargo, Madureira, 2008). These data were analyzed by means of an artificial neural network and statistically by multiple regression analysis. The data reinforce the importance of the description of voice quality based on the phonetic model and the relevant role of keysepeech segments.

Key words: voice quality; auditory perception, phonetics, evaluation; corpus

Introduction

The present study aims at evaluating a corpus designed for voice quality evaluation, based on the phonetic description of voice quality model (Laver, 1980), addressing the role of keysepeech segments in detecting voice quality settings in the Brazilian Portuguese context. The *Vocal Profile Analysis Scheme- VPAS*, which is based on a phonetically grounded description of voice quality (Laver, 1980), is the result of the continuous work of a team of researchers from *Queen Margaret University College- Edinburgh* (Laver et al, 1981; Laver, 2000; Laver, Mackenzie-Beck, 2007).

The application of the VPAS at the Brazilian context to investigating linguistic, paralinguistic and extralinguistic uses of vocal quality and acoustic and physiological correlates of settings of voice quality besides the experience derived from structuring a voice quality database, from introducing the model in workshops and from answering questions, posed by learners in these workshops, about voice data collection procedures and application procedures, led a group of researchers at LIAAC- PUCSP (Integrated Acoustic Analysis and Cognition Laboratory of the Catholic University of São Paulo) to systematize and improve the material aimed at instructional purposes (Camargo, Madureira, 2008).

The phonetic approach to voice quality, based on the susceptibility of speech segments to the effects of voice quality settings was considered in order to elaborate a corpus of 10 keysepeech sentences (Camargo, Madureira,

2008) specially designed for auditory voice quality evaluation by VPAS, as an adaptation for Brazilian context (BP-VPAS).

This adaptation of the VPAS to the Brazilian Portuguese context followed a comprehensive theoretical critical review of the bases of the model profile. The adaptation of the terminology took into account advances in speech science research related to studies on the physiology and speech signal research (Hammarberg, Gauffin, 1995; Chasaide, Gobl, 2005).

Methods

The corpus was recorded by 60 speakers and retrieved in the voice database, being 12 men and 48 women averaged 39,7 years-old. (19 to 58). It contained repetitions of 10 keyspeech sentences (ks) built with keyspeech segments (named as *ks fala*, *ks cidadeSP*, *ks Roberta*, *ks Lara*, *ks Liliane*, *ks garoto*, *ks Cbispos*, *ks detest*, *ks Santos*, and *ks filosofia*) and semi-spontaneous speech samples as answers to the request: "Talk about the city where you were born". These samples were recorded in a radio studio and analysed by means of perceptual (vocal profiles in terms of vocal tract, tension and phonatory settings) and acoustic (f0 measures- median, quartil semi-amplitude and 99,5% quantil- extracted automatically by means of Expression Evaluator script- Barbosa, 2009- version 2011) procedures.

These data were analyzed by means of an artificial neural network and statistically by multiple regression analysis, in order to investigate the speech samples (keyspeech sentences) as related to the evaluation of voice quality settings (in scalar degrees between 1 and 6).

Results

The results contemplate some correlations, detected by means of statistical analysis, between perceived voice quality settings and characteristics of the keyspeech sentences of the corpus. The degree of voice quality settings was found to be related to their incidence in the distinct sentences evaluated.

Almost all the keyspeech sentences were suitable for analyzing both phonatory and non-phonatory voice quality settings evaluated with degree 3 or higher (Figure 1): tongue tip/blade (advanced); jaw (minimized range), tension (laryngeal hyperfunction) and phonatory (creaky and harsh voice).

For the purpose of differentiating degrees lower than 3, the keyspeech sentences were useful to detect distinct lip (spreading, labiodentalization and minimized range), jaw (open), tongue tip/blade (retracted), tongue body (fronted, backed, lowered and minimized range), velopharyngeal (nasal), pharynx-constriction/expansion and tension (vocal tract hyperfunction) voice quality settings.

The semi-spontaneous speech samples allowed the identification of lip spreading, and labiodentalization, advanced and lowered tongue body, nasal, lowered larynx, and breathy voice quality settings (Figure 1).

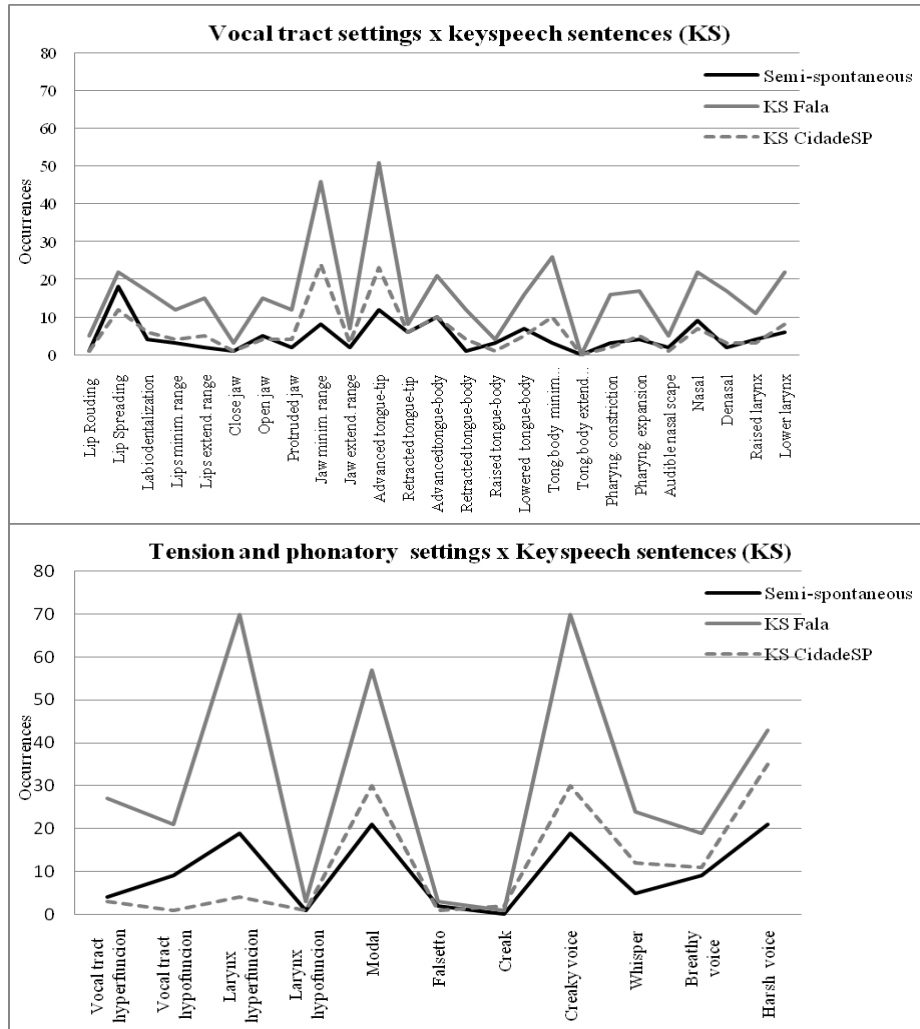


Figure 1. Perceptual evaluation of vocal tract, tension and phonatory settings based on keyspeech sentences (ks) and semi-spontaneous corpus.

For the group of f_0 measures, it was found that median values of f_0 of ks fala were related to that found in semi-spontaneous recordings. However, the f_0 values of all the other keyspeech sentences turn out to be higher. The quartile amplitude were similar between semi-spontaneous samples and ks

detesto, *ks santos* and *ks filosofia*. The semi-spontaneous samples and *ks Roberta* were found to be similar in relation to the 99,5% quantil.

The *ks fala* (*O objeto de estudos da fonética é esta complexa, variável e poderosa face sonora da linguagem: a fala*) allowed the identification of a wide range of voice quality settings, probably due to its composition. It contained as key segments all oral vowel sounds, three nasal vowels, fricatives, plosives, liquids and nasals). Furthermore it presented no difficulties to be produced.

Conclusions

The data reinforce the importance of the description of voice quality based on the phonetic model and the relevant role of some keyspeech segments in the corpus designed for voice quality evaluation.

Acknowledgements

We acknowledge Plínio Barbosa from UNICAMP for SG Expression Evaluator Script- 2011 Version to the database. This study has been supported by Research and Teaching Comission grant at Pontifical Catholic University of São Paulo (CEPE-PUCSP: 09/639).

References

- Barbosa, P. A. 2009. Detecting changes in speech expressiveness in participants of a radio program In Proc. of Interspeech. v.1,2155-2158. Brighton, Reino Unido.
- Camargo, Z., Madureira, S. 2008. Voice quality analysis from a phonetic perspective: Voice Profile Analysis Scheme Profile for Brazilian Portuguese (BP VPAS). In Proc.of the 4th Conf on Speech Prosody v.1, 57–60. Campinas, Brazil.
- Chasaide,N.A.; Gobl, C. 2005. On the relation between phonatory quality and affect. In Hardcastle, W.J., Mackenzie-Beck, J.(eds). A figure of speech: a festschrift for John Laver. Mahwah: Lawrence Erlbrum, 323-346.
- Hammarberg, B., Gauffin, J. 1995. Perceptual and acoustics characteristics of quality differences in pathological voices as related to physiological aspects. In Fujimura, O. and Hirano, M (eds), 1995, Vocal fold physiology, 283-303. San Diego, Singular Publishing.
- Laver, J. 1980. The phonetic description of voice quality. Cambridge University Press. Reprinted.2009.
- Laver, J. 2000. Phonetic evaluation of voice quality. In: Kent, R.D., Ball, M.J. (eds) The handbook of voice quality measurement. San Diego, CA: Singular publications.
- Laver, J., Wirs, S., Mackenzie, J., Hiller, S.M. 1981. A perceptual protocol for the analysis of vocal profiles. Edinburg University Department of Linguistics Work in Progress 14,139-155.
- Laver, J; Mackenzie-Beck J. 2007. Vocal Profile Analysis Scheme -VPAS. Queen Margareth University College-QMUC, Speech Science Research Centre, Edinburgh.

Dynamic differences in child bilinguals' production of diphthongs

Vincent Chanethom

Department of Linguistics, New York University, USA

Abstract

This paper examines the cross-linguistic phonetic interactions in the production of diphthongs by French-English bilingual children. The study aims to investigate (1) whether English diphthongs (e.g. /aɪ/ as in *bye*) and French tautosyllabic vowel-glide combinations (e.g. /aj/ as in *baille* 'yawn') have different phonetic implementations, and if so, (2) whether bilingual children maintain two separate categories. Using SSANOVA to compare formant contours, the results showed distinct phonetic patterns for each category. The results also indicated that bilingual children with a reduced amount of input in French show overlapping acoustic properties.

Key words: Bilingualism, diphthongs, SSANOVA, language acquisition.

Introduction

The current study examines the phonetic interaction between the languages of French-English bilingual children with respect to diphthong production. Previous studies (e.g. Flege 2002) indicated that early bilinguals show smaller degrees of L1-L2 interference than late bilinguals. The majority of these studies, however, investigated the phonetic properties of child bilingual speech using what Tsukada et al. (2005) called a "retrospective developmental design" (p.266), in which adult bilinguals with different ages of first exposure to L2 were compared. Furthermore, a large number of these studies centered mainly on the acquisition of monophthongs. Few studies have examined whether complex segments, such as diphthongs, follow the same pattern of acquisition. For instance, tautosyllabic vowel-glide (VG) combinations in English (e.g. *bye*) and in French (e.g. *baille* 'yawn') have different phonological status. This VG combination corresponds to a single segment (i.e. a diphthong) in English, but two separate segments (i.e. vowel+glide) in French. This paper presents a picture-naming experiment with (1) adult monolingual speakers to examine whether these two categories also have different phonetic implementations, and (2) 6-7 year old bilingual children to investigate whether or not they maintain two separate categories.

To compare these two categories, the current study uses a relatively novel methodology, the Smoothing Spline ANOVA (Davidson 2006). This is a better statistical tool for analyzing diphthongs than previously used techniques. With this technique, it is not only possible to determine whether there are significant acoustic differences in formant structure, but also where these differences lie. Harrington and Cassidy (1994) argue that the best way

to represent diphthongs is by capturing their dynamicity, hence the need of multiple reference points for their analysis. For instance, Tsukada (2008) took three measurements, at the 20th, 50th, and 80th percentiles along the formant trajectories. However, three points are not sufficient to fully capture differences between diphthongs.

In this study, 50 equidistant reference points were extracted using a Praat script (Boersma and Weenink 2011). The curves formed by these 50 points were then submitted to statistical comparisons using SSANOVA, which was run in R. In order to locate the region(s) along the curves with significant differences, 95%-Bayesian confidence intervals were constructed around the main effect smoothing splines (see Figure 1). The parts of the curves where the 95%-confidence intervals for each main effect smoothing spline do not overlap indicate the regions where the two curves are significantly different.

The experiment

Participants

The experiment involved two sets of participants in order to address two specific questions. First, to determine whether English diphthongs (e.g. /aɪ/) and French tautosyllabic VG sequences (e.g. /aj/) differ at the phonetic level, 6 adult subjects participated in this experiment: 3 native speakers of English and 3 native speakers of French. The English native speakers were recorded in the US, while the French native speakers were recorded in France.

A second set of participants comprised of bilingual children was also recruited to investigate whether these children maintain separate categories for the diphthongs or whether they equate them. Four French-English bilingual children between the ages of 6 and 7 years old living in the US participated in the experiment (BH, ZM, IP, ZC). Both IP and ZC attend a bilingual school where French and English are used as languages of instruction. BH and ZM, on the other hand, attend English-only schools. None had any history of speech delay, developmental disability, or neurological impairment.

Stimuli and procedures

The current study is part of a larger project, which involves four English diphthongs (/aɪ, aʊ, eɪ, oʊ/) and three French VG tautosyllabic sequences (/aj, ej, uj/) in varying phonetic environments. Using a picture-naming task, multiple repetitions of the stimuli were elicited from each adult monolingual participant in his or her native language and from each bilingual child in both languages. Due to space limitation, this paper only examines a subset of the acquired data, including the English /aɪ/ (e.g. *bye*) and /eɪ/ (e.g. *obey*) diphthongs in comparison to the French VG sequences /aj/ (e.g. *baille* ‘yawn’) and /ɛj/ (e.g. *abeille* ‘bee’) respectively, following the voiced bilabial stop /b/. Five repetitions of each word from each participant were

analyzed for the study. A total number of 140 tokens (20 per bilingual child and 10 per adult monolingual) were included in the analysis. Fifty-point formant contours (F1 and F2) were obtained from each token and submitted to statistical analyses using SSANOVA for the cross-linguistic comparisons of the word-pairs *baille* vs. *bye* and *abeille* vs. *obey*.

Results

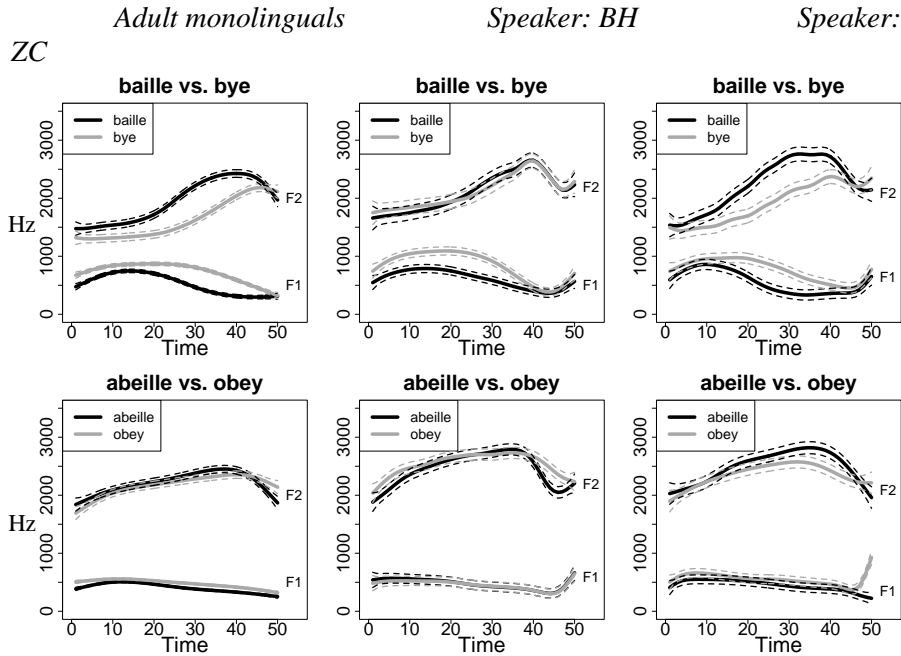


Figure 1. Main effect smoothing spline estimates (solid lines) for cross-linguistic word-pairs with their respective 95%-confidence intervals (dotted lines) for both F1 and F2.

As illustrated in Figure 1, the adult monolinguals' acoustic implementations of the French and English categories are significantly different for the pair /aj/-/aɪ/ but not /ɛj/-/eɪ/, as indicated by the overlapping confidence intervals for the *abeille* vs. *obey* comparison. One major difference between *baille* and *bye* for the adult groups is the presence of an off-glide on F2 trajectories for the French /aj/ after reaching its target. This off-glide does not occur on the English F2 pattern, which may be an indication that the French category involves an independent glide.

Only two bilingual children are represented in Figure 1, because ZM and IP showed similar patterns as BH and ZC respectively. BH shows overlapping phonetic properties for both /aj/-/aɪ/ and /ɛj/-/eɪ/ pairs. On the other hand, ZC patterns the same way as the adult monolinguals with distinct

phonetic implementations for /aj/ and /aɪ/. As for the /ɛj/-/eɪ/ pair, ZC only shows a slight difference between the two categories, in which the target position of /ɛj/ appeared to be more front than that of /eɪ/.

Conclusions

This study indicated that French-English bilingual children differ in diphthong production as a function of amount of input in each language. While the children who attend bilingual schools managed to maintain separate categories for the English diphthong /aɪ/ and French VG sequence /aj/, the ones who attend English-only schools produced these categories with overlapping acoustic properties. Although no differences were found for the adults, the slight differences between /ɛj/ and /eɪ/ for speaker ZC may be attributable to hypercorrection aimed at maximally separating her two languages.

Acknowledgements

I would like to thank the following people for their input on this work: Lisa Davidson, Maria Grigos, Jon Brennan, and Sean Martin.

References

- Boersma, P., Weenink, D. 2011. Praat: doing phonetics by computer [Computer program]. Version 5.2.22, retrieved 14 April 2011 from <http://www.praat.org/>
- Davidson, L. 2006. Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *Journal of the Acoustical Society of America* 120, 407-415.
- Flege, J. 2002. Interactions between the native and second language phonetic systems. In Burmeister, P., Piske, T. and Rohde, A. (eds.) 2002. *An Integrated View of Language Development: Papers in Honor of Henning Wode*, 217-244. Wissenschaftlicher Verlag Trier: Trier.
- Harrington, J., Cassidy, S. 1994. Dynamic and target theories of vowel classification: evidence from monophthongs and diphthongs in Australian English. *Language and Speech* 37, 357-373.
- Tsukada, K. 2008. An acoustic comparison of English monophthongs and diphthongs produced by Australian and Thai speakers. *English World-Wide* 29, 194-211.
- Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., Flege, J. 2005. A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics* 33, 263-290.

Processing Russian inflectional morphology: A PET study of verb generation

Tatiana Chernigovskaya¹, Kira Gor², Galina Kataeva³, Alexander Korotkov³, Maxim Kireev³, Kristina Memetova³, Svyatoslav Medvedev³

¹St. Petersburg State University, Russia

²University of Maryland, USA

³The Institute of the Human Brain, Russian Academy of Sciences

Abstract

For many years, acquisition of verbal morphology has been an issue of much debate within linguistic theory, research on first and second language acquisition, mental grammar representation, and cognitive processing. The proponents of the Dual Mechanism account claim that regularly inflected forms are computed by a rule-processing system, while morphologically irregular forms are processed in associative memory. The opposite Single Mechanism account claims that both regular and irregular forms are processed in associative memory. In the present PET study significant rCBF increase in the anterior cingulate gyrus of the left hemisphere (24 BA) was shown for regular verb generation. This may support the rule-based (de)compositional account.

Key words: inflectional morphology, anterior cingulate cortex, error detection system.

Introduction

For many years, acquisition of verbal morphology has been an issue of much debate within linguistic theory, research on first and second language acquisition, mental grammar representation, and cognitive processing. The “English past tense debate”, which addresses the issue of “nature vs. nurture”, explores whether the processing of regular and irregular morphology is driven by two distinct mechanisms, or by one single mechanism. The proponents of the Dual Mechanism account claim that regular forms are computed by a rule-processing system, while irregular morphological forms are processed in associative memory. The opposite view claims that both regular and irregular forms are processed in associative memory (see for discussion Chernigovskaya & Gor, 2001; Gor 2010). Jaeger et al. (1996) has made an attempt to find the cerebral localization of regular vs. irregular English verb processing, with the data discussed in support of the Dual Mechanism account. However, Chandler & Skousen (1997) published their own interpretation showing that Jaeger’s results are not in conflict with the Single Mechanism Account.

The available brain-imaging data on regular/irregular verb generation are conflicting, which reflects the general controversies in the state-of-the-art neurolinguistic data interpretation in (e.g., Demonet et al. 2005, Sidtis 2006).

This research suggests that being one of the basic brain mechanisms, the error-detection processor (Bechtereva et al., 1968, 2005) may be implicated in linguistic procedures – as it does with all the other cognitive tasks – from routine actions to arithmetic problem solving. Its function is to compare the current status with virtual templates and rules stored in memory. Following it a certain processing routine is chosen under appropriate circumstances.

Our approach predicts that if a verbal form is generated on the basis of a default template (regular forms) it should be automatically checked for rule adequacy, and thus the error detector will be activated. Conversely, if there is no difference between rule-based procedures and associative memory retrieval, its activation will take place in both cases. As the key area for this mechanism is anterior cingulate cortex (Mathalon et al., 2003), it was selected as the region of interest that will make it possible to distinguish between the two accounts.

The aim of the study was to reveal the level of functional activity in anterior cingulate cortex caused by the generation of regular/irregular verbs, nonce verbs, and plural of nouns.

Methods

Subjects

Twelve native right-handed Russian-speaking volunteers, age 22-29, provided written informed consent for participation in the study.

Experimental task

Stimuli (4 word lists in black letters on a white background) were presented at random on a color monitor 1,5 m from the nasion. The stimuli were: the i-class semi-regular verbs (warm-up), singular nouns (N), irregular verbs (I), the aj-class regular verbs (R), the aj-class nonce verbs (Q). During each of the four experimental conditions, subjects viewed the sequence of 65 words at a rate of one word every 2 s, with exposure time from 80 to 160 ms with 1850 ms intervals. Subjects produced 1st person singular past tense forms for verbs and plural forms for nouns. Participants were asked to perform each task continuously for approximately 135 s, because a minimum of 90 s of stimulus data duration is required for acquiring a statistically valid image of brain activity.

Data acquisition and analysis

During PET scans the regional cerebral blood flow (rCBF) was measured with Scanditronix PC2048-15B PET-scanner; the standard protocol was described elsewhere (Kataeva & Korotkov, 2007). PET data processing was performed with SPM8 and WFU Picatlas software (Maldjian et al., 2003). After spatial normalization of images, the mean rCBF was measured in four volumes of interest (VOIs) corresponding to Brodman areas (BA) 24 and 32 of left and right hemispheres. Differences in global flow were discarded by

normalization for the mean acquired activity in the field of view (Arndt et al., 1996) and rCBF was expressed as percentages of mean global CBF. A repeated measures MANOVA was used for statistical analysis with the test conditions as within-subjects effect. The significance level was set at $p=0.01$. The post-hoc analyses with $p=0.01$ (Fisher LSD test) were used to assess the individual differences.

Results and discussion

The evaluation of rCBF found that the PET-condition with regular verbs generation was characterized by increased levels of regional cerebral blood flow in comparison with all the other conditions, within the volume of interest corresponding to 24 Brodmann area (BA) in the anterior cingulate gyrus of the left hemisphere: $R>N$ ($p<0,0004$); $R>Q$ ($p<0,0001$) and $R>I$ ($p<0.0009$).

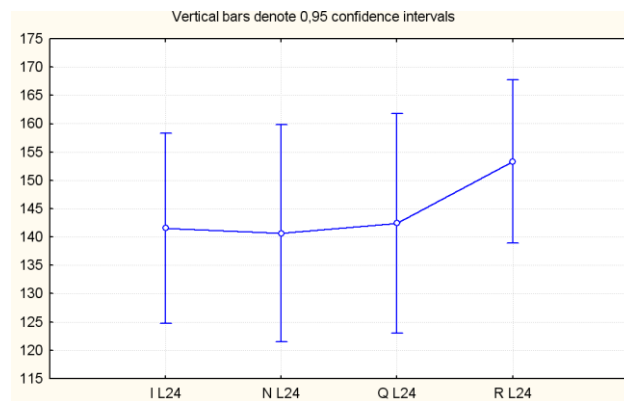


Figure 1. Generation of regular verbs associated with increased rCBF in left anterior cingulated cortex (BA 24). Notations: R- regular verb production; I- irregular verb production; N- noun production; Q – nonce verb production; L 24 – left BA 24.

A significant rCBF increase in the anterior cingulate gyrus of the left hemisphere (24 BA) was shown for regular verb generation in comparison with all the other conditions (irregular, nonce verbs, and nouns). These data indicate that left anterior cingulate cortex is an important brain area subserving the mechanism of regular verbal processing. This is consistent with the rule-based (de)compositional account: The error detector checks if the word matches the systematic and productive rules; it is not activated when the inflected forms are idiosyncratic. It might be argued that these findings are compatible with both the Dual and Single Mechanism accounts. And indeed, BA 24 activation cannot exclude the possibility of a unitary rule-based morphological mechanism for inflectional paradigms. Event-

related fMRI imaging is a promising method to explore the relationship between regular and irregular verbal processing.

Acknowledgements

Grants from the Russian Ministry of Education and Science (# 16.740.11.0113, 02.740.11.0369), RFRH (# 10-04-00056a; 10-06-00726a), OBN RAS 2011.

References

- Arndt, S., Cizadlo, T., O'Leary, D., Gold, S., Andreasen, N.C. 1996. Normalizing counts and cerebral blood flow intensity in functional imaging studies of the human brain. *NeuroImage*, 3, 175-184.
- Bechtereva, N.P., Gretchin, V.B. 1968. Physiological foundations of mental activity. *Int. Rev. Neurobiol.*, 11, 329-352
- Bechtereva, N.P., Shemyakina, N.V., Starchenko, M.G., Danko, S.G., Medvedev, S.V. 2005. Error detectors mechanisms of the brain: background and prospects. *Intern. J. of Psychophysiology*, 58, 227-234.
- Démonet, J.F., Thierry, G., Cardebat, D. 2005. Renewal of the neurophysiology of language: functional neuroimaging. *Physiol Rev.*, 85, 49-95.
- Chandler, S., Skousen, R. 1997. Analogical modeling and the English past tense: a reply to Jaeger et. al. 1996 <<http://humanities.byu.edu/am/>>
- Chernigovskaya, T., Gor, K. 2001. The Complexity of Paradigm and Input Frequencies in Native and Second Language Verbal Processing: Evidence from Russian. *Language and Language Behavior*, 3 (II), 20-37.
- Gor, K. 2010. Beyond the obvious: Do second language learners process inflectional morphology? *Language Learning*, 60.1, 1-20.
- Jaeger, J., Lockwood, D., Kemmerer, R., Van Valin, R., Murphy, R., Kahlak, H. 1996. A positron emission tomography study of regular and irregular verb morphology in English. *Language*. 72. 451–497.
- Kataeva, G.V. and Korotkov, A.D. 2007 The regional cerebral blood flow pattern of the normal human brain and its factor structure. *Human Physiology*, 33.4, 383-387
- Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fmri data sets. *NeuroImage*, 19, 1233– 1239.
- Mathalon, D.H., Whitfield, S.L., Ford, J.M. 2003. Anatomy of an error: ERP and fMRI. *Biolog. Psychol.*, 64, 119-141.
- Sidtis D. 2006. Does functional neuroimaging solve the questions of neurolinguistics? *Brain Lang.*, 98, 276-90.

The terminal contour of Italian semi-spontaneous instructions

Domenico Di Russo

Università di Roma La Sapienza, Italy

Abstract

This article aims to analyze the intonative features of instructions in semi-spontaneous Italian, with particular regard to their terminal contour, in order to understand which is its semantic importance and, from this point of view, what is the nature of relation between intonation and modality.

Index terms: instruction, intonation, modality, semantics, terminal contour.

Introduction

It is well known that the notion of *terminal contour*, i.e. the part of the intonative profile that goes from the last tonic syllable to the end of utterance, is the touchstone of studies about relations between *intonation* and *modality*, so that the distinctive and intuitively normative character of one-to-one correspondence, or rather coalescence, between few basic terminal contour prototypes and few basic sentence modalities is now generally accepted, even by the most important Italian studies on intonation (Chapallaz 1964; Lepschy 1978: 133; Canepari 1985; De Dominicis 1992; Bertinetto & Magno Caldognetto 1993; Sorianello 2006: 118-28).

The case of instructions

In accordance both with the notion of *utterance* as a *sense unit* (Benveniste 1964: 122) that, traced in relation to an enonciative coordinates system (Culioli 1978: 129), displays all cooperating elements which play together to the construction of its signification in continuous, unitary and undivisible *unicum* of its intonative profile, and with the notion of *modality* as an active operation of the speaker's rational activity (Bally 1932: par. 28), in the light of the formal semantics studies tradition inspired by the equally important logic tradition (Lyons 1977, Palmer 1986, Grice 1990), we consider here a particular modality which belongs to one of the five modal classes that we may recognize in spoken language through a tridimensional representative system with an objective (*alethic* modalities), a subjective (*epistemic*, *appreciative* and *volitive* modalities) and an inter-subjective axis (*deontic* modalities) (Le Querler 1996: 63-67), i.e. the case of 528 *instructions* taken from a sample of 15 semi-spontaneous Italian *Map Task* dialogues recorded in 15 representative Italian cities and selected from the dialogical section of *CLIPS corpus* (<http://www.clips.unina.it>). The instructions status consists in giving directions to co-speaker concerning an object, an event or a situation,

they correspond to 81% of all 652 deontic modalities and 13,5% of all 3898 modal utterances and are shared out as follows: 57 in the dialogue of Bari (Ba), 37 in that one of Bergamo (Bg), 38 in Cagliari (Ca), 43 in Catanzaro (Cz), 45 in Florence (Fi), 48 in Genoa (Ge), 28 in Lecce (Le), 26 in Milan (Mi), 11 in Naples (Na), 24 in Palermo (Pa), 29 in Parma (Pr), 31 in Perugia (Pg), 17 in Rome (Rm), 53 in Turin (To), 41 in Venice (Ve) (Di Russo 2011).

As we can note considering, for example, the most simple terminal contour profiles and the ones made up by the mix of these simple tunes, terminal contours configuration shows a very great variety (Table 1).

Table 1. The terminal contour of instructions

Terminal contour types	Instructions number	Percentage (%)
fall	42	8,0
mid-level	3	0,6
rise	23	4,4
fall-low mid level	13	2,5
fall-rise	145	27,5
mid level-fall	5	0,9
mid level-rise	14	2,7
rise-fall	13	2,5
rise-high mid level	1	0,2
fall-low mid level-fall	1	0,2
fall-low mid level-rise	3	0,6
fall-rise-fall	26	4,9
fall-rise-high mid level	12	2,3
mid-level-fall-low mid level	1	0,2
mid-level-fall-rise	4	0,8
mid-level-rise-fall	1	0,2
rise-fall-low mid level	1	0,2
rise-fall-rise	17	3,2
rise-high mid level-fall	1	0,2
rise-high mid level-rise	2	0,3
others	200	37,9
Total	528	100

If we look at terminal contour types distribution from a diatopical point of view, dwelling above all on the most large ones, we may see that their wide diversity is equally shared by all different dialogues instructions. In descending order, in fact, we can note that the 145 *fall-rising* types are so distributed: 6 in Ba, 11 in Bg, 9 in Ca, 4 in Cz, 14 in Fi, 19 in Ge, 7 in Le, 4

in Mi, 4 in Na, 9 in Pa, 15 in Pr, 9 in Pg, 6 in Rm, 23 in To, 5 in Ve. Among the 42 *falling* types of terminal contour, we can find out: 7 in Ba, 4 in Bg, 4 in Cz, 1 in Ge, 6 in Le, 5 in Mi, 1 in Pa, 3 in Pr, 3 in Pg, 1 in Rm, 8 in Ve. The 26 *fall-rise-falling* types are so shared out: 8 in Ba, 6 in Cz, 1 in Fi, 1 in Ge, 2 in Le, 2 in Mi, in Pr, 3 in Pg, 1 in Ve. As regards the 23 *rising* types, we have: 3 in Ba, 5 in Ca, 1 in Cz, 1 in Fi, 1 in Ge, 3 in Mi, 2 in Pg, 1 in Rm, 2 in To, 4 in Ve. Concerning the 17 *rise-fall-rising* types, we can notice this situation: 4 in Ba, 2 in Ca, 1 in Cz, 1 in Fi, 2 in Ge, 2 in Le, 2 in Mi, 1 in Pg, 1 in To, 1 in Ve.

Regarding the distribution of the 14 *mid level-rising* ones: 2 in Ba, 4 in Bg, 1 in Ca, 1 in Ge, 1 in Mi, 1 in Pa, 1 in Rm, 2 in To, 1 in Ve; for the 13 *fall-low mid level* types, there are: 1 in Ca, 4 in Cz, 2 in Fi, 1 in Pa, 3 in Pg, 2 in Ve; for the 13 *rise-falling* types: 1 in Ba, 1 in Ca, 1 in Cz, 1 in Fi, 1 in Le, 1 in Pa, 1 in Pg, 2 in Rm, 3 in To, 1 in Ve; for the 12 *fall-rise-high mid level* ones: 2 in Ba, 6 in Cz, 2 in Fi, 1 in Ge, 1 in Pg; for the 5 *mid level-falling* ones: 1 in Cz, 1 in Ge, 1 in Pa, 1 in Pr, 1 in Ve; for the 4 *mid-level-fall-rising* ones: 2 in Mi, 2 in To; for the 3 *fall-low mid level-rising* ones: 1 in Pa, 1 in Pr, 1 in To; for the 3 *mid-level* ones: 1 in Fi, 1 in Pg, 1 in Rm; for the 2 *rise-high mid level-rising* ones: 1 in Ba, 1 in Pr. Then, we find a *rise-high mid level* type in Pg; a *fall-low mid level-falling* one in Le; a *mid-level-fall-low mid level* one in Ve; a *mid-level-rise-falling* one in Ge; a *rise-fall-low mid level* one in Ba; and a *rise-high mid level-falling* one in To.

Finally, we may count at least 200 instructions with other very different and more complex kinds of terminal contour and, if we then examine the part of intonative profile that goes from the last turning point, i.e. the last point of f_0 variation movement, to the end of the utterance, we can find two groups of utterances: the former is constituted by 185 falling movements, the latter by 343 rising movements, both scattered in so several different terminal contours that what emerges is an equally heterogeneous outline.

Conclusion

So, if we pay attention to those results, we may reach the conclusion that each of the many possible different kinds of terminal contour is capable of meaning, and concretely means, the same deontic value of instruction. Hence we may reasonably conclude that: 1) the terminal contour, like other intonative elements, proves to be not the primary term of modal comparison but only the part of a more complex prosodic game that shows all the limits of current representative models of intonation (Martin 2005, 2009: 69-83) which too often confuse linguistic phenomena dynamics with their representations; 2) each terminal contour type, according to the whole enonciative coordinates system, can mean any modal value and, more generally, make any sense, (Bolinger 1986: 13); 3) the potentially unlimited

semantic plasticity of intonation is just the evidence of its arbitrary character, which proves its status as an integral part of the linguistic system (Saussure 1916; Bally 1932: par. 39, 44; Fónagy 1987: 82).

References

- Bally, C. 1932. *Linguistique générale et linguistique française*. Leroux, Paris.
- Benveniste, É. 1966. Les niveaux de l'analyse linguistique. In Benveniste, É., *Problèmes de linguistique générale*, 119-131. Gallimard, Paris.
- Bertinetto, P.M., Magno Caldognetto, E. 1993. Ritmo e intonazione, in Sobrero, A. (a cura di) *Introduzione all'italiano contemporaneo*, 2 voll., Laterza, Roma-Bari, vol. I, 141-92.
- Bolinger, D. 1986. Intonation and emotion, Contribution to the Round Table «Phonetics and emotion», in «Quaderni semiotici», 1, 13-21.
- Canepari, L. 1985. *L'intonazione*. Linguistica e paralinguistica, Liguori, Napoli.
- Chapallaz, M., 1964. Notes on the Intonation of Questions in Italian. In Abercrombie, D., Fry, D., MacCarthy, P.A.D., Scott, J.L., Trim, J.L.M. (eds), *In Honour of Daniel Jones*, 306-312. Longmans, London.
- Culioli, A., Valeurs aspectuelles et opérations énonciatives: l'aoristique, 1978, in Id., *Pour une linguistique de l'énonciation*, 3 voll., Ophrys, Paris, 1990-99, vol. II, 127-43.
- De Dominicis, A., 1992. *Intonazione e contesto*. Uno studio su alcuni aspetti del discorso in contesto e delle sue manifestazioni intonative, Edizioni dell'Orso, Alessandria.
- Di Russo, D., 2011. *Intonazione e modalità nella significazione*, Dottorato di Ricerca in Filologia, Linguistica e Letteratura, Università di Roma La Sapienza.
- Fónagy, I., 1987. Vocal Expression of Emotions and Attitudes, in «VS. Quaderni di studi semiotici» 47-48, 65-85.
- Grize, J.-B., 1990. *Logique et langage*, Ophrys, Paris.
- Karcevskij, S., 1931. Sur la Phonologie de la Phrase, 1931, in «Travaux du Cercle Linguistique de Prague», vol. IV, 188-227.
- Le Querler, N., 1996. *Typologie des modalités*, Presses Univ. de Caen, Caen.
- Lepschy, G.C., 1978. Appunti sull'intonazione italiana, in Id., *Saggi di linguistica italiana*, Il Mulino, Bologna, 127-42.
- Lyons, J., 1997. *Semantics*, 2 vol., Cambridge University Press.
- Martin, P., 2009. *Intonation du français*. Paris, Armand Colin.
- Id., *Modelli di analisi e sistemi di etichettatura prosodica*, AISV, 2005, Salerno.
- Palmer, F., 1986. *Mood and Modality*, Cambridge University Press.
- Saussure, F. de, 1916. *Cours de linguistique générale*. Paris, Payot.
- Sorianello, P., 2006. *Prosodia. Modelli e ricerca empirica*. Roma, Carocci.

Models or strategies? On the perception of ambiguous words

Anzhelika Dubasova

Laboratory for Cognitive Studies, St. Petersburg State University, Russia

Abstract

Models of lexical ambiguity resolution differ mainly with respect to the role played by context in lexical access. In my paper I focus on two opposing views: those of autonomous access models (non-context-oriented) and selective access models (context-oriented). I analyse two types of situations of: a. the activation of a wrong meaning of an ambiguous word unpredictable from both models and b. humour based on puns. Proceeding from my analysis I propose to include “negative” factors (those leading to errors) and consider existing models as strategies, consequently, treating unsuccessful lexical ambiguity resolution as the influence of “negative” factors or as a selection of the wrong strategy.

Key words: lexical ambiguity, ambiguity resolution, meaning selection

Introduction

Although most words in any language have more than one meaning, we seem to have little difficulty understanding texts containing ambiguous words. It means that lexical ambiguity problems that we might experience get solved somehow. The question “how exactly?” encouraged a number of eye-tracking experiments (Duffy et al. (1988); Rayner & Frazier (1989); Pacht & Rayner (1993); Sereno et al. (2006); etc.) elucidating variables that affect the resolution of lexical ambiguity. The next stage in these studies was to develop theory about the mechanisms involved in ambiguity resolution. These theories differ mainly with respect to the role played by context in lexical access. I will focus on the two opposing views: those of autonomous access and selective access models (see e.g. Reichle et al. (2006); Vu et al. (2000); Connine, Blasko, & Wang (1994); Till, Mross, & Kintsch (1988)).

The current models might reflect general mechanisms of ambiguous words’ perception. However, it begs a question: since there are a number of lexical ambiguity resolution theories, doesn’t it suggest the reference of a particular model to its own part of linguistic reality? To validate this suggestion we should demonstrate situations of lexical ambiguity perception being at variance with the assumptions of either of the models.

a. The activation of a wrong meaning unpredictable from either of the models

Sample 1 [N 1] (translated from Russian, brackets and italics added).

“an examination period... I was standing on the platform at 4 a.m. and waiting for a train. Here it comes. The conductor says something about

‘билеты’ [*tickets* | examination questions]. Without thought, I take my credit test book and offer it to her wondering : where are the ‘билеты’ [*examination questions* | tickets]?”

The Russian word “билет” is a biased ambiguous word with the dominant meaning of ‘a document used to gain admission to a location or event; a passage ticket’ (its frequency of occurrence is about 76% [N 2]). The meaning of ‘examination card; examination questions’ is one of the 3 subordinate meanings of the word, its frequency is about 4%.

The situation (a railway station, a train, a conductor) unambiguously points to the dominant meaning of the word acting as a resolving context.

Neither autonomous access models nor selective access models can explain the activation of a wrong meaning here. The autonomous access models posit that all word meanings are exhaustively accessed at a rate that is proportional to their frequency, irrespective of context. Thus, the dominant (here, correct) meaning was to be activated in Sample 1. The selective access models posit that only the contextually appropriate meanings of ambiguous words are accessed at a rate that is proportional to their frequency (again it is the correct meaning that was to be activated).

I propose the following explanation of the error. Word meaning activation hardly occurs with all the other word meanings deactivated. Some meanings might have been activated before the actual process of word perception started. These create a “background” context for information processing. As a rule, the perception of ambiguous words is determined by resolving context and meaning frequencies and the “background” context does not influence word processing considerably. But in some cases the “background” context (thinking about exams in Sample 1) becomes stronger suppressing the “figure” (thoughts about a trip).

Samples of this type demonstrate a contradiction between the assumptions of the models and reality. However, the contradiction is smoothed over by including certain factors preventing the access of the appropriate meaning of an ambiguous word into the model. Let us treat this case as a competition between “positive” (ensuring successful perception) and “negative” (causing errors) factors. The “negative” factors can be roughly described as the above-mentioned “background” context, leading to a *two*-level competition, where the first is between the meanings of the ambiguous word (see e.g. Reichle et al. (2006)), and the second is between the factors affecting meaning selection (and thus, affecting previous level competition).

b. Humour based on puns

Sample 2 (translated from Russian, brackets and italics added).

“- Your breakfast, sir!

- Well, bacon, eggs, beans... and where is the *toast* [‘тост’]?

- To your health, sir!”

An ambiguous word “тост” is also a biased one with the dominant meaning ‘ceremonial act of drinking in honour of someone or something’ (its frequency is about 96%). It is important that, despite its high frequency, a listener /reader does not retrieve the dominant meaning (with the comical effect provided by unexpectedness). The prior context causes the retrieval of the subordinate meaning ‘bread browned with dry heat’ (its frequency is about 4%). The example supports selective access models.

Sample 3 (translated from Russian, brackets and italics added).

“There is a board in our cafeteria where anyone can write his/her criticism of our cooking. Today there was an entry “the beans are too ‘острый’ [*spicy* | sharp]”. I couldn’t refrain from adding “and pricked my finger”.

The Russian word ‘острый’ has 10 meanings, among them ‘sharp’ (with frequency about 13%) and ‘spicy’ (with frequency about 3%). This example supports autonomous access models.

Natural communication situations generate evidence for both opposing views. For this reason I suggest treating the existing models as strategies: “context-oriented vs. non-context-oriented”, each responsible for its own “domain”. Hence, we deal here with a *three*-level competition: a. between meanings of an ambiguous word, b. between the factors affecting meaning selection, and c. between the mentioned strategies (so that level C affects level B and level B affects level A).

Summary

Taking into account just “positive” factors would be idealistic, to an extent as they assume a listener / a reader free from any other preoccupation at the moment. Considering the “negative” factors would increase the explanatory power of these models, although treating models as strategies seems preferable here, since it allows the introduction of a concise model covering greater number of situation types. The suggested innovation could also explain misunderstanding as an effect of “negative” factors or a selection of a wrong strategy.

Notes

1. Most samples are from <http://bash.org.ru/> (collection of quotes from ICQ, IRC, forums and everyday life); about 100 relevant samples were analysed.
2. Relative frequencies were assessed on the basis of a corpus search (Russian National Corpus <http://ruscorpora.ru>).

Acknowledgements

The research is supported by grants from Ministry of Education and Science of the Russian Federation (state contracts NN 16.740.11.0113, 02.740.11.0369), Russian Scientific Foundation for Research in Humanities (N 10-04-00056a).

References

- Connine, C.M., Blasko, D., Wang, J. 1994. Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception & Psychophysics* 56, 624–636.
- Duffy, S.A., Morris, R.K., Rayner, K. 1988. Lexical ambiguity and fixation times in reading. *Journal of Memory and Language* 27, 429–446.
- Pacht, J.M., Rayner, K. 1993. The processing of homophonic homographs during reading: Evidence from eye movement studies. *Journal of Psycholinguistic Research* 22, 251–271.
- Rayner, K., Frazier, L. 1989. Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15, 779–790.
- Reichle, E.D., Pollatsek, A., Rayner, K. 2006. Modeling the effects of lexical ambiguity on eye movements during reading. B.R.P. Van Gompel, M.F. Fischer, W.S. Murray, R.L. Hill (Ред.), *Eye movements: A window on mind and brain* (pp. 271–292). Oxford, England: Elsevier.
- Sereno, S.C., O'Donnell, P.J., Rayner, K. 2006. Eye movements and lexical ambiguity resolution: Investigating the subordinate bias effect. *Journal of Experimental Psychology: Human Perception and Performance* 32, 335–350.
- Till, R.E., Mross, E.F., Kintsch, W. 1988. Time course of priming for associate and inference words in a discourse context. *Journal of Verbal Learning and Verbal Behavior* 16, 283–298.
- Vu, H., Kellas, G., Metcalf, K., Herman, R. 2000. The influence of global discourse on lexical ambiguity resolution. *Memory & Cognition* 28, 236–252.

Spectral properties of fricatives: a forensic approach

Natalie Fecher

Department of Language and Linguistic Science, University of York, York, UK

Abstract

This paper reports on the acoustic-phonetic analysis of the voiceless fricatives /s, ʃ, f, θ/ taken from high-quality recordings of six native British English speakers reading phonetically controlled stimuli under various face disguise conditions. Speech samples were extracted from an audio-visual ‘face cover’ corpus that was collected for the purpose of investigating multimodal speech and speaker recognition in a forensic context. Findings are discussed with regard to constraints in speech production and acoustic damping effects caused by certain mask materials.

Key words: forensic phonetics, acoustics, fricatives, spectral moments

Introduction

Forensic speech scientists are particularly confronted with the inherent intra- and inter-speaker variation in speech, language and the human voice. The practitioners, however, have to cope not only with this speaker-induced variability, but also with the discrepancies between studio-quality speech material and often low-quality (authentic) forensic samples. One aspect that has to date not been studied in detail is the extent to which changes in a speaker’s visual appearance have an impact on the speech signal. Hereby we refer to various types of headgear and face-concealing garments (henceforth FCGs) typically worn for occupational, recreational and religious purposes, or for the commission of crimes such as armed robberies and assaults. All FCGs obscure parts of the talker’s articulators and will, to a varying degree, get in the way of their normal functioning and are likely to cause further modifications to the acoustic (and consequently the auditory) signal.

Material and method

Six speakers (3 M, 3 F, aged 21-36) were recorded in a professional TV studio at the University of York. All were native British English speakers with normal vision and hearing, training in phonetics, and no experience in regularly wearing any kind of face covering. Their task was to repeatedly read aloud a list of 64 phonotactically legal /C₁aC₂/ syllables embedded in the carrier phrase *He said <stimulus>*, each time wearing one of the FCGs shown in Figure 1. For the acoustic analysis, two tokens per syllable position per fricatives /s, ʃ, f, θ/ were extracted. On the basis of relevant previous studies (e.g. Jongman et al., 2000) five parameters capturing spectral properties and intensity of the frication noise were chosen for analysis. A repeated-measures ANOVA was applied to investigate the effects of the

dependent factors *intensity*, *peak frequency*, and the first four statistical moments of the averaged FFT power spectrum, i.e. *centre of gravity*, *variance*, *skewness* and *kurtosis*, on the independent factors *place of articulation* (POA), *syllable position* (SYL) and *disguise condition* (FCG). Measures were taken from spectra computed over non-filtered/non-pre-emphasised speech (48kHz/16-bit) in *Praat 5.1.44* (see also Fecher, 2011).



Figure 1: Control condition and face coverings. Selection criteria for the FCGs were forensic relevance, mask material and parts of the face covered.

Results

The ANOVA revealed a high level of significance ($p < .001$) for the main effects of POA ($F(3,12)=63.53$), SYL ($F(1,4)=69.48$) and FCG ($F(7,28)=12.68$), and also for the interactions between POA and SYL ($F(3,12)=29.00$) and POA and FCG ($F(21,84)=3.33$), on the *intensity* of all fricatives.

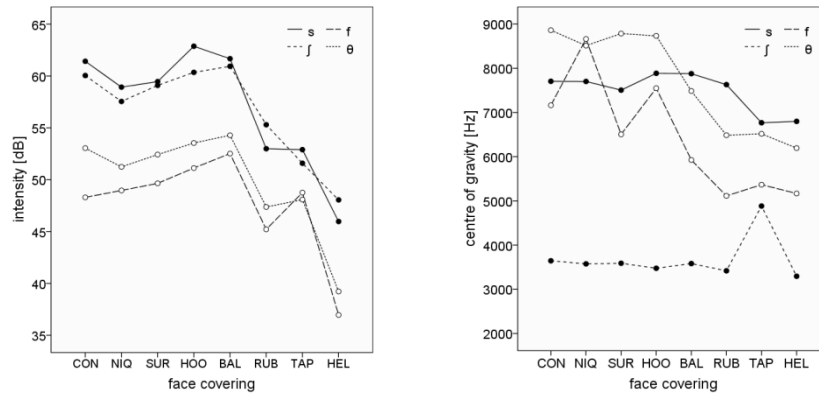


Figure 2 a) Intensity (*left*) and b) centre of gravity (*right*) for all fricatives and FCGs, averaged across speakers and syllable positions.

As expected, certain mask materials (HEL/RUB/TAP) absorb the sound energy more than others (see Figure 2a). However, intensity measures for several face coverings (BAL/HOO/SUR/NIQ) show even higher values than the control. This may have been caused by some speakers compensating for the (perceived) decreased loudness of their speech by speaking with greater vocal effort and thus ‘overriding’ the acoustic damping effects.

The main effects of POA ($F(3,12)=43.92$), SYL ($F(1,4)=125.93$) and FCG ($F(7,28)=8.81$) on *centre of gravity* were also found to be significant at $p<.001$. As can be seen in Figure 2b, certain masks lower the centre of gravity for /f, θ/, but not for /s, ʃ/. These FCGs are likely to absorb energy particularly in higher frequency bands, and the non-sibilants may be more prone to this damping due to their greater spectral diffuseness and overall lower energy (Shadle&Mair, 1996). The same explanation holds for the *variance* (standard deviation) of non-sibilants, which is generally higher for the non-sibilants than for the sibilants (POA: ($F(3,12)=49.21$, $p<.001$).

Regarding the *peak frequency* there is a significant main effect at $p<.001$ for POA ($F(3,12)=14.18$) and at $p<.01$ for SYL ($F(1,4)=30.92$), but none for FCG ($F(7,28)=1.03$, $p=.43$). *Skewness* and *kurtosis* were significant at $p<.001$ only for POA ($F(3,12)=82.22$; $F(3,12)=32.33$) (see Figure 3).

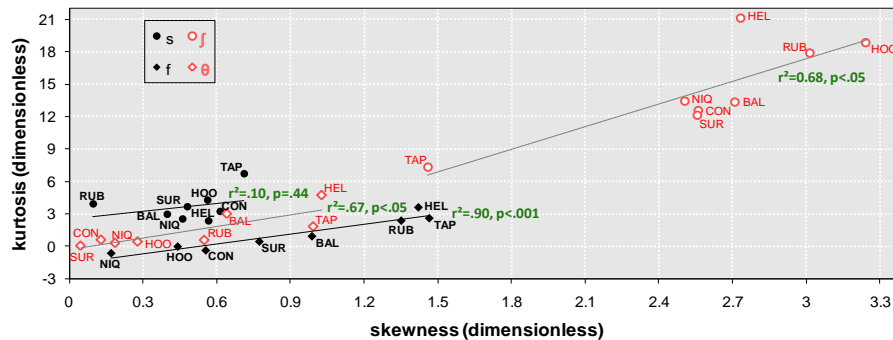


Figure 3: Skewness and kurtosis for all fricatives. The means for both spectral moments determine the data point for each face covering. For each fricative except /s/ a significant positive correlation can be observed.

Discussion and conclusion

The choice of fricatives for this study was motivated by their high perceptual confusability, their relevance as consonantal features in forensic phonetics, and an anticipated larger attenuation by certain FCGs of energy in higher frequency bands that are particularly discriminative for this phoneme class. The shifts in the spectral patterns may be caused by acoustic damping effects of certain mask materials, leading to energy being absorbed at higher frequencies (Llamas et al., 2009; Watt et al., 2010). When an FCG obstructs the talker's face, it may also interfere with speech production. Physiological and somatosensory effects, such as lip/nose contact, restricted jaw movement (Iskarous et al., 2009) and skin stretching (Fuchs et al., 2010) can lead to modified articulatory behaviour. Simultaneously, each subject may reveal individual compensation strategies, e.g., by increasing the overall intensity (loudness). The speakers may produce speech with greater pulmonary/glottal

effort, with the effect that the spectrum is not amplified uniformly, but higher frequencies are elevated more (Sluijter et al., 1997). This effect may be reinforced by some masks covering the speakers' ears, in this way impairing auditory self-monitoring. The outcome of the present study will be of particular value in combination with the results of upcoming perception tests in which participants will be asked to identify the stimuli under visually and acoustically degraded conditions. This may hinder the successful mapping between facial cues and auditory percepts, reducing intelligibility. Findings will be beneficial for forensic work (e.g., speaker comparison, transcription tasks, earwitness testimony) and research on audio-visual speech processing.

Acknowledgement

This work is supported by the FP7 Marie Curie ITN "Bayesian Biometrics for Forensics" (<http://bbfor2.net>). Many thanks to Dom Watt for his support.

References

- Fecher, N. 2011. Speaking under cover: The effect of face-concealing garments on spectral properties of fricatives. Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong, August 2011 (accepted).
- Fuchs, S., Weirich, M., Kroos, C., Fecher, N., Pape, D., Koppetsch, S. 2010. Time for a shave? Does facial hair interfere with visual speech intelligibility? In: Fuchs, S., Hoole, P., Mooshammer, C., Zygis, M. (eds.). Between the regular and the particular in speech and language. Frankfurt/M.: Peter Lang, 247-264.
- Iskarous, K., Shadle, C., Proctor, M. 2008. Evidence for the dynamic nature of fricative production: American English /s/. Proceedings of the 8th International Seminar on Speech Production, Strasbourg, 405-408.
- Jongman, A., Wayland, R., Wong, S. 2000. Acoustic characteristics of English fricatives. *J. Acoust. Soc. Am.* **108** (3), 1252-63.
- Llamas, C., Harrison, P., Donnelly, D., Watt, D. 2009. Effects of different types of face coverings on speech acoustics and intelligibility. *York Papers in Linguistics (Series 2)* **9**, 80-104.
- Shadle, C., Mair, S.J. 1996. Quantifying spectral characteristics of fricatives. Proceedings of Interspeech, Philadelphia, 1521-24.
- Sluijter, A. M. C., van Heuven, V. J., Pacilly, J. J. A. 1997. Spectral balance as a cue in the perception of linguistic stress. *J. Acoust. Soc. Am.* **101** (1), 503-513.
- Watt, D., Llamas, C., Harrison, P. 2010. Differences in perceived sound quality between speech recordings filtered using transmission loss spectra of selected fabrics. Talk at the IAFPA Conference 2010, Trier.

Vocabulary learning strategies among advanced Turkish learners

Ioannis Galantomos

Department of Mediterranean Studies, University of the Aegean, Greece

Abstract

The goal of this paper is to examine the vocabulary learning strategies used by Turkish learners at an advanced level. Current research suggests that vocabulary is an important component of language proficiency. Nevertheless, for a long period of time it was a neglected aspect during foreign language teaching and learning. Vocabulary learning strategies play a crucial role in vocabulary acquisition. In order to investigate the vocabulary learning strategies used by advanced Turkish learners of Greek origin, we conducted a small-scale experiment. Our results indicate a preference for strategies of formal vocabulary learning and practicing over spontaneous vocabulary learning strategies and self-initiated independent vocabulary learning.

Key words: foreign language learning, Turkish as a foreign language, vocabulary learning strategies

Introduction

During the past decades, vocabulary has been a neglected aspect in the field of foreign language learning (Meara 1980). Nevertheless, current research has demonstrated the important role vocabulary plays in the acquisition of a foreign language (e.g. Meara 2002). Moreover, the above mentioned neglect opposes to the importance attributed to vocabulary by learners themselves (Meara 1980).

One of the aspects that have attracted research attention is the study of the strategies that learners use in order to cope with the meanings of new words. Early research into learners' strategies began in the 1970s and by the 1990s the interest has moved to vocabulary learning strategies (Schmitt 2010). Experiments carried out so far reveal that learners use many strategies when they encounter unknown words. This finding is due to the fact that learners find it more feasible and manageable to deal with individual lexical items rather than with larger units (Schmitt 2010).

The increasing interest in vocabulary learning strategies is also reflected in the various taxonomies that have been proposed in the relevant literature (for a review cf. Pavičić Takač 2008). For instance, Schmitt (1997) identified four major groups of strategies, namely, social (=cooperation with other learners), cognitive (=language manipulation or transformation), metacognitive (control of the learning process) and memory (=relating a new word with existing knowledge) strategies. Later, this classification was enriched with determination strategies (=discovering a new word's meaning

without the involvement of other learners or speakers). Another influential taxonomy is the one proposed by Nation (2001) and according to which three general classes of strategies are introduced, namely, planning (choosing what to focus on and when to focus on), sources (=finding clues about new words) and processes (=establishing knowledge).

In view of the above, we carried out a small-scale experiment in order to investigate the vocabulary learning strategies used by advanced Turkish learners.

Method

Participants

Our study consisted of 22 tertiary Greek students who learn Turkish as a foreign language in the Department of Mediterranean Studies at the University of the Aegean (Rhodes island campus). Regarding our participants' sex, 4 were males and 18 females. Their mid age was 22.7 years old and they were all attending regular Turkish classes at an advanced level (3rd level according to department's terminology).

Instrument

The data on vocabulary learning strategies were collected through adopting and adapting an existing questionnaire consisting of 27 vocabulary learning strategies (the questionnaire is available upon request). This questionnaire has already been used in a similar survey (Pavičić Takač 2008). These 27 strategies are categorized into three main groups, namely, strategies of formal vocabulary learning and practicing, self-initiated independent vocabulary learning and spontaneous vocabulary learning. According to this scheme, strategies 1-11 fall within the first group, strategies 12-20 fall within the second group and the remaining ones (i.e. 21-27) within the third group. Each strategy is followed by a three-type scale (i.e. 1: never, 2: sometimes and 3: always).

Procedure

The questionnaire was administered to all students enrolled on advanced Turkish classes in January 2011. The researcher, after being introduced by the instructor, explained the goal of the study and provided detailed instructions on how to fill out the above mentioned questionnaire. The whole procedure lasted for approximately 25 minutes. The data were analysed by means of SPSS for Windows and the statistical criterion χ^2 ($p < 0.05$) was used to examine the statistical significance of our findings.

Results

Our results indicate a preference for strategies of formal vocabulary learning and practicing over the other two types of strategies. In particular, our

participants used more strategies of the first group (7 out of 11, $p < 0,05$). On the other hand, only 4 and 2 strategies of the second and third group of strategies respectively showed a statistical significance ($p < 0,05$).

Discussion

Our findings (i.e. preference for formal vocabulary learning strategies) can be attributed to the context of learning Turkish. Formal learning strategies in general are employed in a classroom-based setting. Turkish is learnt by our participants as a foreign (and not as a second) language, that is, there is no real and actual need for communication in the target language outside classroom. According to Pavičić Takač, (2008), reliance on formal vocabulary learning strategies involves instrumental motivation and is oriented towards formal language learning tasks. For instance, our participants' reports that they test themselves (strategy 4), make lists of new words (strategy 9) and translate new lexical items into their mother tongue (strategy 11) are related to the fact that they expect to be tested and evaluated for their knowledge of Turkish language by their instructors in the form of, let's say, a final exam. Therefore, the fact that they showed a preference for formal vocabulary learning strategies over more functional (/communicative) strategies is due to the requirements and norms of instructed language learning.

Experimental data support our finding. Thus, Lamb (2004) suggests that the degree of exposure to the target language can have an impact on the development of more or less and specific vocabulary learning strategies. In addition, Pavičić Takač (2008) reports that the major difference observed between the two groups which participated in her study is related to the large amount of language input.

Moreover, the preference for formal vocabulary learning strategies may be related to the sex factor. The majority of our sample consisted of females. Research data (e.g. Jiménez Catalán 2003) suggest that females rely more often on formal rule strategies than males. This finding is corroborated by Gu (2002) who claims that females tend to use more strategies that contribute to success in EFL learning. Nevertheless, the limited sample of our survey and the lack of equivalence between the two sexes do not allow us to reach final conclusions regarding this aspect.

Conclusions

To sum up, the results of our survey reveal a preference of our subjects for formal vocabulary learning strategies over more functional ones. This finding is attributed to the classroom-based context of learning the target language and it can probably be related to sex differences, although more experimental data are required in order to support such a claim.

References

- Gu, Y. 2002. Gender, Academic Major, and Vocabulary Learning Strategies of Chinese Efl Learners. *RELC Journal*, 33, 35-54.
- Jiménez Catalán, R. 2003. Sex differences in L2 vocabulary learning strategies. *International Journal of Applied Linguistics*, 13, 54-77.
- Lamb, M. 2004. It depends on the students themselves': Independent language learning at an Indonesian state school. *Language, Culture and Curriculum*, 17, 229-245.
- Meara, P. 1980. Vocabulary Acquisition: A Neglected Aspect of Language Learning. *Language Teaching and Linguistics: Abstracts*, 13, 221-246.
- Meara, P. 2002. The rediscovery of vocabulary. *Second Language Research*, 18, 393-407.
- Nation, I.S.P. 2001. *Learning Vocabulary in Another Language*. Cambridge: CUP.
- Pavičić Takač, V. 2008. *Vocabulary Learning Strategies and Foreign Language Acquisition*. Clevedon: Multilingual Matters.
- Schmitt, N. 1997. Vocabulary learning strategies. In Schmitt, N. and McCarthy, M. (eds.) 1997, *Vocabulary: Description, Acquisition and Pedagogy*, 198-227. Cambridge: CUP.
- Schmitt, N. 2010. *Researching Vocabulary. A Vocabulary Research Manual*. London: Palgrave Macmillan.

Vowel-colour associations in non-synesthetes: A study with Spanish and Arabic participants

Pilar Mompeán Guillamón

English Department, University of Murcia, Spain

Abstract

The present paper aims at contributing to the field of sound symbolism and, more specifically, to the association between sounds and colours as established by non-synesthetes. A study based on a forced-choice task performed by Spanish and Arabic speakers is presented. The study asks participants to listen to primary cardinal vowels and choose from a range of colours the one considered most suitable for the sound. The data gathered reinforce previous results that non-synesthetic participants are able to significantly associate vowel sounds and colours at a better than chance degree. However, results seem to go against the general idea that the associations are cross-linguistic, although the phenomenon itself seems to be.

Key words: cardinal vowels, colours, synesthesia, synesthetic symbolism

Introduction

Research on synesthetic sound symbolism has shown that participants from the normal population associate sounds and properties of objects at a better than chance degree (Sapir 1929; Westbury 2005). On the other hand, research on synesthesia has shown that synesthetes are able to automatically associate two senses, namely hearing and vision (Simner 2006; Ward & Mattingley 2006).

Taking into account the results obtained in the aforementioned areas, the question that emerges is: Will non-synesthetes be able to establish significant connections between two domains typically attributed to the synesthetic population, i.e. between sounds and colours? The research carried out in this respect (Miyahara *et al.* 2006; Wrembel 2007; Wrembel & Rataj 2008) shows that Japanese, Polish and English speakers are able to significantly relate vowel sounds and colours.

In an attempt to provide more evidence for this field and expand the range of languages studied, an experiment was designed. Native speakers of Spanish and Arabic had to listen to vowel sounds and choose the colour suitable for the sound listened to. Results show that participants significantly associated sounds and colours, although differences were also found between the two languages studied, probably due to linguistic and cultural factors.

Experiment

The present study contributes to the field of the association between sounds and colours by non-synesthetes. Two research questions were proposed: a) Will non-synesthete participants significantly associate sounds and colours?; b) Will the associations established be common to the two languages? Based on the results obtained in previous research, it was hypothesized that: i) Participants would be able to establish significant associations between sounds and colours; ii) The associations would be different for the two languages studied, although some common elements could be found. These differences would be basically due to two types of factors: linguistic and cultural.

Participants

For the experiment with Spanish participants, the data of 52 native speakers were taken into account. For the experiment with Arabic participants, the data of 48 native speakers of Darija were considered.

Stimuli

Two types of stimuli were used, auditory and visual, and these were the same for the two languages studied, so that, differences were found, these were due to the different stimuli used.

Auditory stimuli were the eight primary cardinal vowels (Jones [1918]1967). Cardinal vowels are not real vowels, but reference points for all the languages of the world. The following cardinal vowels were used: 1 /i/, 2 /e/, 3 /ɛ/, 4 /a/, 5 /ɑ/, 6 /ɔ/, 7 /o/, 8 /u/.

Visual stimuli were 12 coloured squares situated against a light grey background. Eleven squares were designed following Berlin & Kay's (1969) findings of the basic colour terms for all the languages of the world (white, brown, purple, red, grey, yellow, pink, green, orange, black and blue). A twelfth square was added that participants could use in case they felt that none of the colours corresponded to the sound listened to.

Materials

For Spanish speakers, the materials used were a desktop computer with a 15-inch screen, headphones and a dimmed quiet room. For Arabic speakers, a laptop computer was used, since the experiment had to be carried out in different places. In any case, the environment was always dimmed and quiet.

Procedure

The procedure followed was the same for both languages and the experiment was run by means of a computer programme designed for the purpose of the study. First participants were given an informed consent form and the experiment was explained orally. Then they sat in front of a computer and the programme asked them to choose their native language and provided

them with written explanations. Once they had read the explanations, they were asked about some personal data and required to do a colour-blindness test. If they passed this test, the experiment started automatically. Participants heard one sound and after that a screen with the 12 squares appeared. They had to click on the colour they associated the sound with. The position of the colours changed, so that findings would not be due to location of the colour squares.

Results and discussion

The results obtained in the two experiments were organized and chi-square tests were run in order to check the degree of significance of the associations (see Table 1).

Spanish participants provided significant results. Cardinal vowel 1 was significantly associated with yellow, cardinal vowels 2 and 3 were related to green, cardinal vowels 4 and 5 were connected with red, cardinal vowel 6 was related to pink, cardinal vowel 7 with black, and cardinal vowel 8 was connected with grey and purple.

The results provided by Arabic participants were not as strong as those of Spanish, though still significant. Cardinal vowel 1 was related to grey, cardinal vowel 2 with green and pink, cardinal vowel 3 with green, cardinal vowel 4 with blue and green. Following the tendency of the Spanish, cardinal vowel 5 was related with red, whereas cardinal vowel 6 was associated with black, cardinal vowel 7 with brown and orange, and 8 with pink and red.

Therefore, the results obtained in the three experiments seem to confirm hypothesis 1, that participants would be able to associate sounds and colours at a better than chance degree. These results seem to follow the line of the findings obtained in other works such as Miyahara *et al.* (2006), Wrembel (2007) and Wrembel and Rataj (2008).

Hypothesis 2 is also confirmed, since there are common choices between the languages, but mostly these tend to be language-specific. The differences observed can be due to linguistic and cultural factors. In the case of Spanish, pronunciation and spelling may have conditioned participants' choices; whereas Arabic choices might have been influenced by the small vowel inventory of this language. Culturally speaking, colours may be associated with positive or negative connotations, and this may have conditioned participants' choices. For example, green seems to be of paramount importance for the Arabic culture. This would explain why, at a higher or lower degree depending on the vowel, green is always chosen in Arabic, even for vowels for which in Spanish the choice of green is almost imperceptible.

Table 1. Chi-square results for Spanish and Arabic participants.

Vowels	Spanish	Arabic
/i/	Yellow ($X^2(1)=59.130$, $p<0.005$)	Grey ($X^2(1)=9.824$, $p<0.005$)
/e/	Green ($X^2(1)=42.391$, $p<0.005$)	Green ($X^2(1)=4.888$, $p<0.05$) Pink ($X^2(1)=4.324$, $p<0.05$)
/ɛ/	Green ($X^2(1)=12.057$, $p<0.005$)	Green ($X^2(1)=9.321$, $p<0.005$)
/a/	Red ($X^2(1)=13.684$, $p<0.005$)	Blue ($X^2(1)=4.047$, $p<0.05$) Green ($X^2(1)=4.001$, $p<0.05$)
/a/	Red ($X^2(1)=10.278$, $p<0.005$)	Red ($X^2(1)=8.871$, $p<0.005$)
/ɔ/	Pink ($X^2(1)=9.750$, $p<0.005$)	Black ($X^2(1)=4.047$, $p<0.05$)
/o/	Black ($X^2(1)=15.260$, $p<0.005$)	Brown ($X^2(1)=5.301$, $p<0.025$) Orange ($X^2(1)=5.873$, $p<0.025$)
/u/	Grey ($X^2(1)=8.091$, $p<0.005$) Purple ($X^2(1)=10.582$, $p<0.005$)	Pink ($X^2(1)=4.324$, $p<0.05$) Red ($X^2(1)=6.352$, $p<0.025$)

References

- Jones, D. [1918]1967. *An Outline of English Phonetics*. Cambridge: Heffer.
- Sapir, E. 1929. A study in phonetic symbolism. *Journal of Experimental Psychology* 12, 225-239.
- Simner, J. 2006. Beyond perception: synesthesia as a psycholinguistic phenomenon. *Trends in Cognitive Sciences* 11(1), 23-29.
- Ward, J., Mattingley, J. 2006. Synesthesia: an overview of contemporary findings and controversies. *Cortex* 42, 129-136.
- Westbury, C. 2005. Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language* 93, 10-19.
- Wrembel, M. 2007. "Still sounds like a rainbow" – a proposal for a coloured vowel chart. In *Proc. of the Phonetics Teaching & Learning Language Conference*, August 2007.
- Wrembel, M., Rataj, K. 2008. Sounds like a rainbow. In *Proc. of the 2nd ISCA Workshop of Experimental Linguistics*, August 2008, Athens, Greece, 237-240.

The tip-of-the-tongue phenomenon: Search strategy and resolution during word finding difficulties

Nina Jeanette Hofferberth

Department of German Linguistics, University of Wuppertal, Germany

Abstract

A tip-of-the-tongue (TOT) experience refers to the state in which a speaker is temporally unable to retrieve a word from memory, while being sure that he knows the word. The recovered partial information can consist of competing items that resemble the target word phonologically or semantically and could give rise to competition or conflict during attempts to resolve the TOT. A question that has been discussed recently is whether phonologically similar words block or facilitate lexical retrieval. A study at Frankfurt University (cf. Hofferberth 2008) found that participants preferred searching for their intended word by a semantic search strategy instead of a phonological one.

Key words: speech production, lexical access, word finding failures, blocking, incomplete activation

The tip-of-the-tongue phenomenon

The tip-of-the-tongue (TOT) state has been described as one „in which one cannot quite recall a familiar word but can recall words of similar form and meaning“ (Brown & McNeill 1966: 325). In a TOT state, speakers have a strong feeling of knowing the word, have access to its meaning and (partial) access to its syntactic properties. Speakers cannot retrieve the complete phonological form but are often able to retrieve the first letter, the number of syllables, stress pattern, other letters/phonemes in the word as well as words with similar sound and similar meaning.

Successful lexical retrieval in a TOT state can be imminent (within minutes) or delayed (after some hours or days) and may occur spontaneously (pop-ups). The resolution may be through external search strategies (such as looking up the word in a dictionary or by asking someone) or through internal strategies (such as searching in the alphabet or generating similar words) (cf. Brown 1991).

TOT states may represent the momentary unavailability of an otherwise accessible word (blocking hypothesis) or the weak activation of an otherwise inaccessible word (incomplete activation hypothesis) (cf. Meyer & Bock 1992).

The recovered partial information could give rise to competition during attempts to resolve the TOT. For example, given the cue *Aida Composer*, subjects may recover the initial letter *V*, the final letter *I* and know that it is a short Italian name and could then recall *Vivaldi* instead of *Verdi* (cf. Maril et al. 2001: 658).

Experimental methodology

A multiple single case study with German students at Frankfurt University examined the search strategy to resolve the TOT state (cf. Hofferberth 2008).

Participants

The participants in this experiment were five undergraduates at Frankfurt University who volunteered their services. All subjects were native speakers of German.

Materials

Participants attempted to retrieve 25 rare German target words (18 nouns, 2 verbs, 5 adjectives) from definition, such as “an instrument used for the application of capital punishment by decapitation” for *guillotine*. This study is a replication of the classic questionnaire study by Brown & McNeill (1966). If the target word was known or unknown, the subject could write down the answer on the response sheet or leave it open. If the participant experienced a TOT state, he responded to a series of questions about initial, final and further letter(s) of the word, number of syllables, words with similar sound, words with similar meaning, word class, and any other data that came to mind (e.g., gender of a noun).

Procedure

All participants were tested together in one room. At the beginning of the experiment, each participant received a response sheet. The experimenter explained the tip-of-the-tongue phenomenon and read aloud the printed instructions. Then, the experimenter read the first definition aloud, and participants were allowed 90 sec to fill out the response sheet. At the end of this period, the next definition was presented. This procedure was repeated for each of the 25 words in the list.

Response coding

The evaluation of positive TOTs (word recalled is correct) and unresolved TOTs (no word recalled) can be seen in figure 1. Negative TOTs (word recalled is incorrect) remain unconsidered because the information given by the participants refers to a non-target word.

Results

In 46.4% of the cases, the participants could immediately name the word; in 10.4% they did not know the word; and in 43.2% they experienced a TOT state (14.1% positive TOTs, 11.1% negative TOTs and 64.8% unresolved TOTs). Participants preferred searching for their intended word by a semantic search strategy, i.e. via associations, context, co-hyponyms, such as *Marie Antoinette*, *France*, *bastille*, *axe* to retrieve *guillotine*. The results showed that participants used only marginally the phonological search

strategy (searching in the alphabet *G...* or creating words with similar sound, such as *gelatine* to retrieve *guillotine*).

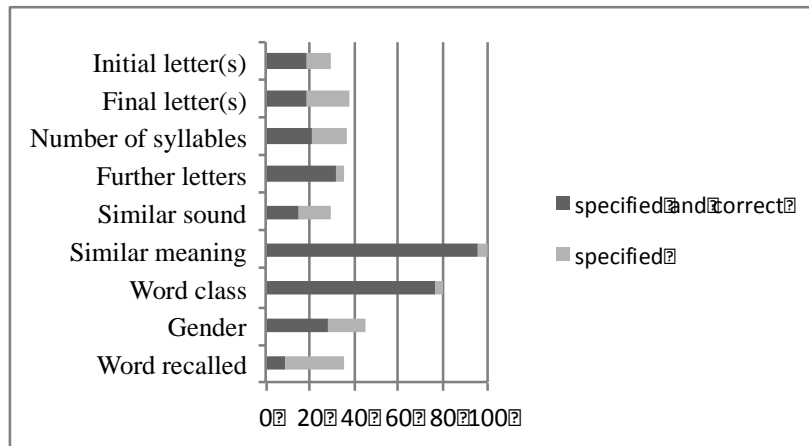


Figure 1. Evaluation of positive TOTs (word recalled is correct) and unresolved TOTs (no word recalled). Main result: All participants experiencing a TOT responded to the category 'words with similar meaning' (100%), 98% of the information given was correct. The preference for resolving a TOT was a semantic search strategy and not a phonological one.

Discussion

The results show that participants used mostly the semantic search strategy by browsing semantic networks (specified in 100%, correct in 98%) and only hardly via phonologically similar words (specified in 29.2%, 50% correct). The preference for the semantic search strategy is more consistent with the blocking hypothesis and could be interpreted in a way that words with similar sound are usually not used to resolve the TOT state because they might block the target.

General discussion and suggestions for future research

TOTs are an important source of information concerning the nature of the processes and architecture of the speech production system. Models of lexical retrieval largely agree on the general distinction between semantic-syntactic (lemma) and word form retrieval processes (lexeme). They differ, however, in the precise architecture of the system. Currently under debate is the relationship between semantic, syntactic, and morpho-phonological units as well as the time-course of their activation.

Experiments by Jones (1989; see also Jones & Langford 1987) and Maylor (1990) found that TOT states were more frequent when target and cue were phonologically related than when they were unrelated or only related in

meaning. These results are more in line with the incomplete activation hypothesis. Meyer & Bock (1992), on the other hand, showed that sound cues were more effective retrieval aids than meaning cues. These results are more consistent with the incomplete activation hypothesis.

The phonological cues used by Meyer & Bock (1992) had the same initial sound and letter, the same number of syllables, and the same stress pattern as the targets but were unrelated to them in meaning (e.g., phonological cue *ignorant* for the target *incubate*). These phonological cues might not block lexical retrieval because they are less related to the target than words with similar rhyme (e.g., the phonological cue *gelatine* would probably block the target *guillotine*). In order to compare different phonological cues (such as the first syllable and the end rhyme) with semantic cues (such as associations and co-hyponyms), reaction time experiments with students at Dusseldorf University will commence in summer in order to examine the preference of the semantic search strategy any further.

References

- Brown, A.S. 1991. A review of the tip-of-the tongue experience. *Psychological Bulletin* 109, 204-223.
- Brown, R., McNeill, D. 1966. The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behaviour* 5, 325-337.
- Hofferberth, N.J. 2008. Das Tip-of-the-Tongue-Phänomen. Eine multiple Einzelfallstudie. Frankfurt University: Unpublished master's thesis.
- Jones, G.V. 1989. Back to Woodworth: Role of interlopers in the tip-of-the-tongue phenomenon. *Memory & Cognition* 17, 69-76.
- Jones, G.V., Langford, S. 1987. Phonological blocking in the tip of the tongue state. *Cognition* 26, 115-122.
- Maril, A., Wagner, A.D., Schacter, D.L. 2001. On the tip of the tongue: An event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron* 31, 653-660.
- Maylor, E.A. 1990. Age, blocking and the tip-of-the-tongue state. *British Journal of Psychology* 81, 123-134.
- Meyer, A.S., Bock, K. 1992. The tip-of-the-tongue phenomenon: Blocking or partial activation? *Memory & Cognition* 20, 715-726.

L2 Greek morphological agreement

Sviatlana Karpava

Department of English, University of Cyprus, Cyprus

Abstract

The present study investigates morphological agreement in L2 Greek verbal and nominal domains by L1 Russian, Bulgarian, and Turkish speakers, with various features such as person, number, gender, and case involved, following Tsimpli et al. 2005. Russian, Bulgarian, and Greek are similar in terms of phi-feature agreement, but different from Turkish. A further distinction will be made for the L2 Greek, between Standard Modern Greek and Cypriot Greek. The analysis of the oral interviews shows that gender, an uninterpretable feature that lacks semantic content (Tsimpli 2004, Franceshina 2001), was the most problematic feature for non-native speakers of Greek, both in the verbal and in the nominal domain. L2 learners either relied on phonological agreement or used the default neuter gender.

Key words: L2 acquisition, morphological agreement, (un)interpretable features, cross-linguistic interference

Introduction

According to the Missing Surface Inflection Hypothesis (Haznedar & Schwartz 1997, Lardiere 1998, Prévost & White 2000), L2 acquisition is UG-constrained, so the operation Agree is available to L2 learners. Morphological agreement is realized post-syntactically at the PF-interface. Syntax and morphology develop independently in L2A without a straightforward mapping between them (Smith & Tsimpli 1995). It is difficult for L2 learners to access and match the morpho-phonological form with the relevant feature(s) (Lardiere 2005). The variability of L2 output by non-native speakers depends on their L1 and the length of exposure and age of onset to L2, the domain of agreement, locality issues, and morphosyntactic features involved.

The present study investigates morphological agreement in L2 Greek verbal and nominal domains by L1 Russian, Bulgarian, and Turkish speakers, with various features such as person, number, gender, and case involved (following Tsimpli et al. 2005). Russian, Bulgarian, and Greek are similar in terms of phi-feature agreement, but different from Turkish. For example, Greek and Bulgarian are +determiner/clitic languages, while Russian and Turkish are –determiner/clitic languages, and Russian, Greek, and Bulgarian are +gender languages, while Turkish is a –gender language.

The study

This study investigates grammatical aspects in the L2 acquisition of Greek, both Standard Modern Greek (SMG) spoken in mainland Greece and

Cypriot Greek (CG) spoken in the Republic of Cyprus. The data come from 40 oral interviews conducted in Greece (with 8 Russian–SMG adult speakers) and Cyprus (with 17 Russian–CG, 8 Bulgarian–CG, and 8 Cypriot Turkish–CG adults). Table 1 shows information on the participants.

The interviews included four parts (Tsimpli et al. 2005): (1) natural conversation on background information, (2) story-telling while describing of eight sets of pictures, (3) two-instruction-giving tasks, and (4) general discussion on various daily life topics.

Table 1. Participants.

Participants	Russian-SMG	Russian-CG	Bulgarian-CG	Turkish-CG
N	8	16	8	8
Gender	5 male 3 female	16 female	5 female 3 male	3 female 5 male
Age mean	31.5	32.8	40.7	64.3
Origin	2 Kazakhstan 2 Georgia 2 Russia 2 Moldavia	6 Russia 4 Belarus 5 Ukraine 1 Moldova	8 Bulgaria	8 Cyprus
Job	6 low 2 high	10 high 6 low	6 low 2 high	2 high 6 low
Education	5 school 3 university	12 university 4 school	4 school/ 4 university	2 university 8 school
LoR mean	13.5	10	2.8	64.3
AoO Mean	19.5	23	37.8	0
MLU	5.05	7.26	4.33	4.27

Results

The analysis of the participants' oral production showed that Bulgarian–CG and Russian–CG speakers omitted more determiners than the other non-native groups. In addition, there were more omissions for neuter and feminine than for masculine determiners. Russian–SMG speakers omitted more feminine determiners than the other participants, while Russian–CG, Bulgarian–CG, and Turkish–CG speakers omitted more neuter determiners. Feminine determiners singular in nominative and accusative case and neuter determiners singular and plural in accusative case had the highest omission rate (Table 2).

Russian–SMG speakers omitted more clitics than the other non-native groups, Bulgarian–CG and Turkish–CG speakers had nearly the same clitic omission rate and performed worse than the Russian–CG speakers. Overall, participants omitted more determiners than clitics. Neutral singular clitics in accusative case had the highest rate of omission (Table 2).

All non-native groups performed more proclisis than enclisis. The participating residents of Cyprus had more enclitics than the participants who live in Greece. This can be explained by the influence of their L2 CG (Table 2). The position of pronominal object clitics in CG is mainly post-verbal, except for certain syntactic environments, such as negation and subjunctive embedded clauses with *na*, which require proclisis, both in CG and SMG. The default clitic placement in SMG is pre-verbal, apart from some syntactic environments, which enforce enclisis, for example, imperatives (Terzi 1999, Grohmann 2011).

All L2 learners had more utterances with null rather than overt subjects, Bulgarian–CG and Turkish–CG speakers had a higher percentage of null subject than the other groups. Turkish is a language with an agglutinating pronominal paradigm and so allows radical pro-drop, while Bulgarian and Russian have fusional or invariant pronominal paradigms and do not allow radical pro-drop; pronouns are not present only in certain contexts (Neeleman & Szendroi 2007).

Table 2. Determiner and clitic omission by L2 groups.

%	Russian-SMG	Russian-CG	Bulgarian-CG	Turkish-CG
Determiner omission	41	53.31	68.8	42.6
Clitic omission	13.5	5.18	11	10.12
Proclisis	5.62	11.68	8.25	15.75

The Russian–CG and Bulgarian–CG non-native groups had more agreement errors than the other non-native groups. Within DP, gender was the most problematic; in subject–verb agreement L2 participants had more errors in number, and subject–adjective mismatches observed were mainly due to gender. Overall, the gender feature was the most problematic in agreement relations for L2 participants (Table 3).

Table 3. Agreement mismatches by L2 groups.

Agreement mismatch %	Russian-SMG	Russian-CG	Bulgarian-CG	Turkish–CG
Features total (41)	4	15	13	9
Gender (23)	3	8	7	5
Number (7)	1	2	3	1
Person (1)	0	0	0	1
Case (10)	0	5	3	2

The analysis of the recordings shows that gender was the most problematic feature for non-native speakers of L2 SMG and CG, in both the verbal and the nominal domain. Grammatical gender is an uninterpretable feature as it does not coincide with natural gender and lacks semantic content (Franceschina 2001, Tsimpli 2004), so L2 learners have problems assigning

gender features and gender concord. L2 learners of either variety of Greek in this study either relied on phonological agreement or used default neuter gender. It has also been found that length of exposure to L2 influences the production of surface morphology (Table 4), in line with findings by Gess & Herschensohn (2001), Lardiere (2005), and Tsimpli et al. (2005).

Table 4. Variables and determiner and clitic omission and agreement mismatches

Variable %	Determiner omission	Clitic omission	Agreement errors
LoR 0–5	66	11	17
LoR 6–10	47	7	9
LoR 11–15	49	6	13
LoR 16–...	42	9	7

References

- Franceschina, F. 2001. Against an L2 morphological deficit as an explanation for the differences between native and non-native grammars. *EUROSLA Yearbook* 1, 143-158.
- Grohmann, K. K. 2011. Some directions for the systematic investigation of the acquisition of Cypriot Greek: A new perspective on production abilities from object clitic placement. In Rinke, E. and Kupisch, T. (eds.), *The Development of Grammar: Language acquisition and diachronic change*. 2011, 179–203.
- Haznedar, B. and Schwartz, B.D. 1997. Are there optional infinitives in child L2 acquisition? In Hughes, Hughes, M., and Greenhill, A. (eds.) *Proc. of the 21st Annual BUCLD*. Somerville, MA: Cascadilla Press, 257-268.
- Lardiere, D. 2005. On morphological competence. In Dekydtspotter, L., Sprouse, R., and Liljestrand, A. (eds.), *Proc. of the 7th GASLA Conference (GASLA 2004)*, Somerville, MA: Cascadilla Press, 178-192.
- Neeleman, Ad., Szendroi, K. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry*, 38, 671-714.
- Prévost, P. and White, L. 2000. Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research* 16, 103-133.
- Smith, N. and Tsimpli, I.M. 1995. *The Mind of a Savant: Language Learning and Modularity*. Oxford: Blackwell.
- Terzi, Arhonto (1999) "Cypriot Greek clitics and their positioning restrictions". In A. Alexiadou, G. Horrocks and M. Stavrou (eds.) *Studies in Greek Syntax*. Kluwer Academic Publishers, pp. 227-240.
- Tsimpli, I.M., Roussou, A., Fotiadou, G., and Dimitrakopoulou, M. in press. The syntax/morphology interface: Agree relations in L1 Slavic/L2 Greek. In Tsoulas, G. and Galani, A. (eds.), *Proc. of the 7th International Conference on Greek Linguistics*, 2005, York, England.
- Tsimpli, I.M., Sorace, A., Heycock, C., Filiaci, F. 2004. First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism* 8, 257-277.

Perceptual and lexical priming of syntactic construction in young children

Meesook Kim

Department of English Language and Literature, Sangji University, Korea

Abstract

Two experiments are reported to investigate the roles of perceptual priming and lexical priming in Korean children's production of passive sentences. Experiment 1 examined whether manipulations of focal attention could affect Korean children's linguistic choices regarding subject role assignment, by using so-called *fish film* (Tomlin, 1997). 5-year-olds and 6-year-olds were not able to produce passives even when a patient was perceptually primed by a flashing arrow. Experiment 2 examined lexical priming effects in children's production of passives. 5-year-olds and 6-year-olds were more likely to use passive sentences if a patient has been verbally primed by the experimenter. These results indicate that young children's production of passive voice is more likely to be affected by linguistic cues than perceptual salience.

Keywords: focal attention, language acquisition, lexical priming, passives, perceptual priming

Introduction

When do speakers make a decision of using passive sentences? A common answer for this question is that speakers may use passives to topicalize the patient in an event, by promoting it to syntactic subject (Dik, 1989; Langacker, 1990). For instance, in the full passive sentence like *The cat was chased by the dog*, the focus is on the cat and what happened to it, rather than on the dog's act of chasing as in the active description of this event. Accordingly, if we could manipulate speakers' attention on the patient, then they would be able to use the passive construction frequently. How can we manipulate speakers/hearers' attention on the patient in an event?

One possibility is that perceptual cues or factors play a crucial role in word order or constituent. Specifically, research on the perceptual salience has shown that perceptual factors like size and color are clearly involved in ordering simple conjoined noun phrases (e.g. A bear and a dog) (Gleitman, Gleitman, Miller & Ostrin, 1996; Osgood & Bock, 1977). In addition, the role of perceptual salience has been empirically tested by Tomlin (1997), by using the so-called *fish film* in on-line production task. Tomlin finds a close relationship between primed referents and subject role assignment. However, the role of perceptual salience in constituent order still remains unclear because some studies find no relationship between initially fixated stimuli and subject role assignment (Griffin & Bock, 2000).

Another possibility is that linguistic cues like lexical or syntactic priming play a crucial role in subject role assignment. For instance, a preceding stimulus like a particular word or sentence increases the likelihood that the speaker will produce an identical word or sentence. In addition, some researchers have argued that in the early stages of syntactic development children have lexically-based word patterns rather than general grammatical form (Olguin & Tomasello, 1993; Tomasello, 2000).

In this paper, two experiments were designed to investigate the role of perceptual salience as well as lexical priming effects in Korean children's production of passive sentences, by using the fish film designed by Tomlin (1997). First, if the experimental manipulation of focal attention affects subject assignment in voice alternation, we expect that it will also be confirmed if tested using data from language acquisition. Therefore, the present study extends Tomlin's (1997) experimental paradigm of attention manipulation to language acquisition, by testing Korean children aged 5 to 6. Another experiment was used to investigate whether linguistic cues like lexical priming would affect children's sentence production in subject role assignment. In this experiment, a preceding stimulus like a lexically-primed referent (either agent or patient) will be heard to each participant. Accordingly, the linguistic Priming tool predicts that Korean-speaking children will be more likely to produce the reference as syntactic subject when they hear the experimenter's sentence.

Experiments

Experiment 1: Perceptual priming paradigm

Participants

Twenty 5-year-old children (age range=5:3-5:11, Mean=5:6) and twenty 6-year-old children (age range=6:2-6:11, Mean=6:8) participated in this experiment.

Materials and Procedure

The experimental materials used in the perceptual priming task were similar to the original fish film designed by Tomlin (1997). However, the present study modified fish film by using a different size of fish, in order to focus on the saliency of the perceptual priming. The experimental framework in the present study consists of a production task where subjects are to produce on-line descriptions of a sequence of 20 events. It shows two computer-animated fish swimming towards each other until they meet and one of them eats the other and swims off the screen.

Results

Table 1. Number and Percentage of Utterances by Each Age Group: Perceptual Priming Paradigm

	Agent-Primed(AP)		Patient-Primed(PP)	
	Active	Passive	Active	Passive
5-year olds	200(100%)	0(0%)	190(95%)	10(5%)
6-year olds	200(100%)	0(0%)	182(91%)	18(9%)

Experiment 2: Lexical priming task

Participants

Twenty 5-year-old children (age range=5:0-5:11, Mean=5:5) and twenty 6-year-old children (age range=6:0-6:10, Mean=6:6) participated in this experiment. They were recruited from daycare centers in the Wonju area.

Materials and procedure

The experimental materials used in the production task were the same as those used in the perceptual paradigm. That is, like Experiment 1, Experiment 2 also used a fish-film which consisted of twenty events. Unlike Experiment 1, the second experiment enforced the child to produce either active or passive sentences by priming a lexical word. For instance, after showing each event of the film to the child, the experimenter asked “what’s happening [What happened] to X” where X was the patient or the agent of a transitive action. Focusing on the child’s attention on the patient by priming a lexical word may create a pressure to produce a passive sentence because the child may be more likely to put the patient in the subject position (Brooks & Tomasello, 1999; Marchman, Bates, Burkardt, & Good, 1991).

Results

Table 2. Number and Percentage of Utterances by Each Age Group: Lexical Priming Paradigm Result

	Agent-Primed(AP)		Patient-Primed(PP)	
	Active	Passive	Active	Passive
5-year olds	199 (99.5%)	1(0.5%)	144 (72%)	56 (28%)
6-year olds	200 (100%)	0 (0%)	115 (57.5%)	85 (42.5%)

Conclusion

Taken together, the results of these two experiments clearly demonstrate that there is a relationship between linguistic priming and children’s production of passive voice, whereas there is no direct relationship between attention and language production. Further investigation will be necessary to examine the role of perceptual salience in children’s sentence production, by manipulating various perceptual cues like colour or size. In addition, further

investigation will be necessary to examine the effects of experimental manipulations of visual attention, by using various verbs for describing visual scenes. Using various verbs for describing an event will allow us to take a closer look at children's developmental patterns of passive voice.

References

- Bock, K. 1986. Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12, 575-586.
- Bock, K. 1990. Structure in language: Creating from in talk. *American Psychologist* 45, 1221-1236.
- Bock, K., Loebell, H., Morey, R. 1992. From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review* 99, 150-171.
- Gleitman, L., Gleitman, H., Miller, C., Ostrin, R. 1996. Similar and similar concepts. *Cognition* 58, 321-376.
- Griffin, Z.M., Bock, J.K. 2000. What the eyes say about speaking. *Psychological Science* 11, 274-279.
- Huttenlocher, J., Vasilyeva, M., Shimpi, P. 2004. Syntactic priming in young Children. *Journal of Memory and Language* 50, 182-195.
- Olguin, R., Tomasello, M. 1993. Twenty-five-month-old children do not have a grammatical category of verb. *Cognitive Development* 8, 245-272.
- Osgood, C., Bock, J. 1977. Salience and sentencing: Some production principles. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory* (pp. 89-140). Hillsdale, N.J.: Erlbaum.
- Tomlin, R. 1997. Mapping conceptual representations into linguistic representations: The role of attention in grammar. In N. Jan & P. Eric (Eds.), *Language and conceptualization* (pp. 162-189). Cambridge: Cambridge University Press.

Time selected multiple algorithms for reliable Fo tracking in difficult recording conditions

Philippe Martin

Univ Paris Diderot, Sorbonne Paris Cité, CLILLAC-ARP EA 3967
UFRL, 175 rue du Chevaleret, 75013 Paris, France.

Abstract

Prosodic data mining in large spontaneous speech corpora often requires acoustic analysis of recordings done in poor conditions. The most detrimental pertain to the absence first harmonic in voiced segments and the presence of echo. In such cases it is very difficult to separate inter vocalic consonantal voicing from echo, or recover Fo from remaining harmonics. To address these problems, 10 different pitch tracking algorithms were implemented in the software program WinPitch, in order to allow an expert user, guided by an underlying displayed narrow band spectrogram, to select manually the appropriate process to deliver a satisfactory Fo curve for a selected time segment.

Key words: intonation, fundamental frequency, spontaneous speech, speech corpus

Introduction

The development of relatively large scale spontaneous speech corpora made possible new investigations in the domain of the interaction between syntax and prosody. With this new kind of data, predictions of well-established theoretical approaches such as autosegmental-metrical can now be confronted with more diverse data than the ones obtained through experimental procedures implying only read or elicited speech.

As most of the pitch curves obtained today use the speech analysis software Praat, problems occur frequently due to the use of a default fundamental frequency tracking algorithm based on a modified autocorrelation method (Boersma, 1993). Not only the frequency scale displayed under Praat lack details and is difficult to read, but the mediocre quality of speech recordings often found in the most interesting cases results in locally erroneous melodic curves, leading at times to misleading theoretical interpretation of the data.

To address this problem, no less than ten different algorithms have been implemented in the analysis software WinPitch, allowing the user to use them selectively on defined time segments, in order to obtain a resulting Fo curve exhibiting a satisfactory match with the first or second harmonic of an underlying displayed narrow band spectrogram. Although this appropriate algorithm manual selection can be time consuming and requires some user expertise in the field, it is possible with this implementation to obtain clean fundamental frequency curves from recordings whose poor quality would otherwise give locally or globally a large number of errors.

Current enemies of Fo tracking

Numerous errors in the Fo tracking process can occur when the hypotheses under which a particular analysis algorithm is designed are not fulfilled. Among the most common due to specific recording conditions we have:

1. Recording microphone with a poor low frequency response preventing the first voice harmonic to be present in the spectrum and making it difficult to recover when the analysed voiced speech segment has few or no harmonics (case of intervocalic stop consonants, of some nasals, etc.);
2. Presence of echo in the recording room, giving trailing frequencies in some part of the spectrum (depending of the room dimensions and wall or window material), easily confused with the first or second harmonic of voiced stops;
3. Use of Automatic Volume Control (AVC) distorting the speech sound dynamic, giving erroneous intensity readings (especially in voiced unvoiced and unvoiced voiced transitions) and possibly saturation the speech signal;
4. When AVC is not used, inappropriate recording input level, giving either a locally saturated speech signal (level to high), or conversely a low signal to noise ration when the input level is too low (distance between the microphone and the speaker to large);
5. Multiple sound sources due to background noise or multiple speaker overlapping segments;
6. Use of low bitrate in mp3, wma or other similar coding. Although the resulting auditory quality of these compression schemes can appear satisfactory, they sometimes give problematic results in case of Fo tracking analysis. The wav format should be used in all cases whenever possible;

Although some of these conditions detrimental to a reliable fundamental frequency analysis can be avoided, a compromise is frequently to be found between the interest of material recorded and its technical qualities. Nevertheless, some simple rules can prevent the data analysis to become an impossible challenge of prosodic analysis.

Multiple Fo tracking strategies

Experts in fundamental frequency tracking have shown for a long time that to date no single algorithm gives reliable results in all cases (Hess, 1983). For instance, in some occasions, an autocorrelation algorithm will give good results for fast Fo change whereas a spectral based analysis will fail, and conversely the autocorrelation will fail in voiced segments correctly processed with a Fourier analysis based algorithm.

To address these limitations, new Fo tracking algorithms have been added to WinPitch in order to offer more flexibility in selecting appropriate algorithms for difficult cases. Presently, these algorithms are: Spectral comb (Martin, 1981), Spectral brush (Martin, 2000), Cepstrum (Noll, 1964), AMDF (Ross et al. 1974), Standard Autocorrelation (Rabiner, 1977), Praat modified autocorrelation (Boersma, 1993) and Yin autocorrelation (2002).

Furthermore, the user can manually 1) define pitch period directly on the speech wave displayed with an appropriate zoom scale; 2) apply the spectral comb method on manually selected specific harmonics and 3) force Fo to be zero for a selected segment. All methods use adjustable parameters, and pitch curves generated elsewhere with the Praat .pitch format can be imported in WinPitch and used as a supplementary method.

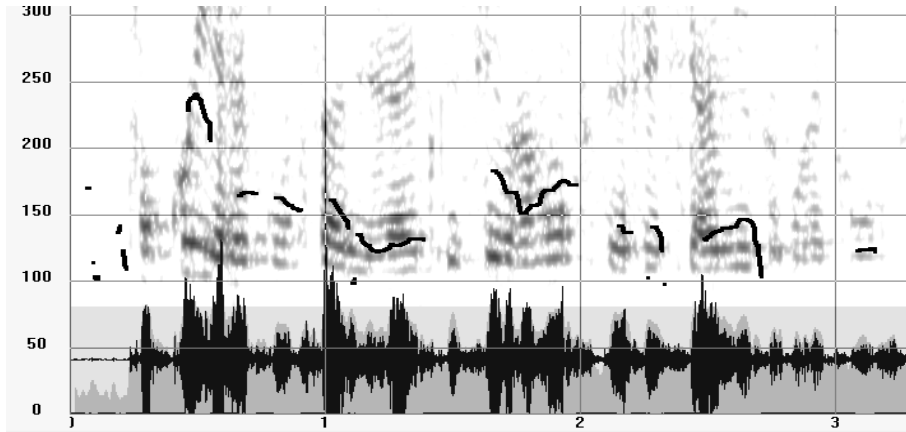


Figure 1. Rough pitch curve obtained by the spectral comb method. In this difficult example, the signal first harmonic is very weak and barely visible on the underlying narrow band spectrogram. Furthermore, the speech signal has been mp3 encoded with a very low bit rate (as used in inexpensive conference speech recorders).

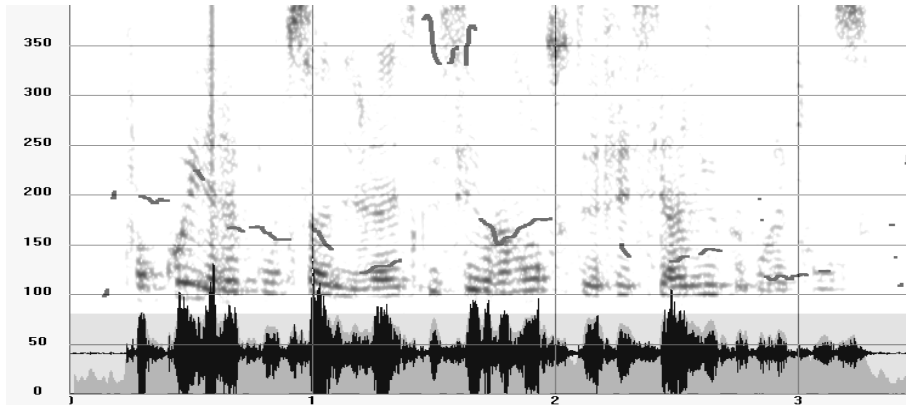


Figure 2. Pitch curve obtained by Praat default autocorrelation algorithm. In some section, the Fo value is either absent or erroneous.

This arrangement was first used on the Rhapsodie project to clean up Fo curves which were mostly tampered with many errors, preventing an

automatic indexing labelling process at a later stage. Figure 3 gives an example of F_0 correction using various algorithms, compared to the commonly used spectral comb (Fig. 1) and Praat pitch curve (Fig. 2).

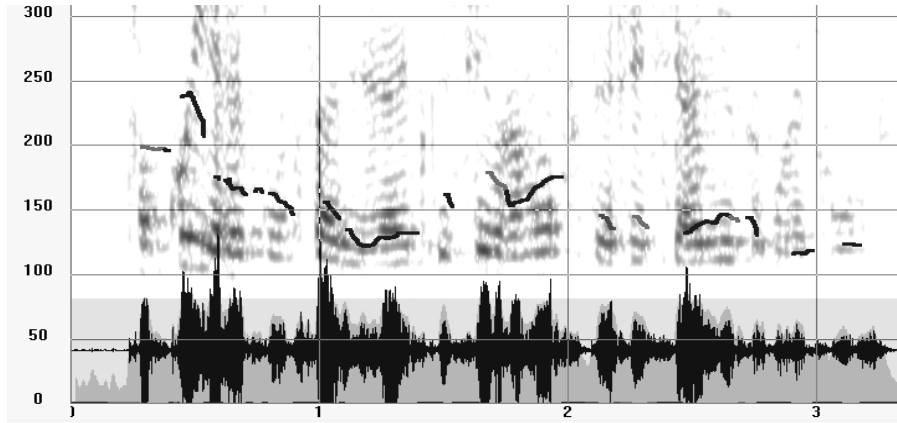


Figure 3. Corrected pitch curve using spectral comb on harmonic selection, AMDF and yin autocorrelation. Sections are displayed in different colours according to the method used locally.

References

- Boersma, P. 1993. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proc. of Institute of Phonetic Sciences of the University of Amsterdam*, vol. 17, 97-110, Amsterdam, The Netherlands.
- de Cheveigné, A., Kawahara, H. 2002. YIN, a fundamental frequency estimator for speech and music. *JASA*, 111:1917-1930.
- Hess, W. 1983. *Pitch Determination of Speech Signals: Algorithms and Devices*. Heidelberg, Germany, Springer-Verlag.
- Martin, Ph. 1981. Mesure de la fréquence fondamentale par intercorrélation avec une fonction peigne. *Proc. 12th JEP, GALF, Montréal*.
- Martin, Ph. 2000. Peigne et brosse pour F_0 : Mesure de la fréquence fondamentale par alignement de spectres séquentiels. *Proc. of XXIIIèmes JEP Aussois, France*.
- Noll, A.M. 1964. Short-time spectrum and cepstrum techniques for vocal-pitch detection. *JASA*, vol. 36, 2, 296-302.
- Rabiner, L.R. 1977. On the use of autocorrelation analysis for pitch detection. *IEEE Trans. Acoust, Speech, Signal Processing*, VOL. ASSP-25, NO. 1, 1977
- Ross, M.J., Shaffer, H.L., Cohen, A., Freudberg, R., Manley, H. J. 1974. Average magnitude difference function pitch extractor. *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, 353-362, Oct. 1974.
- Praat: www.praat.org
- Rhapsodie: Corpus de référence de français parlé, <http://rhapsodie.risc.cnrs.fr/fr/>
- WinPitch: www.winpitch.com

Towards a mental representation of vowel height in SSBE speakers

Kevin Mendousse

School of European Languages & Literatures, University of Auckland, New Zealand

Abstract

Vigorous debate in phonetics and phonology has focused on the cognitive foundation of distinctive feature theory, as well as on the definition of features themselves. In particular, we show that, although the motor-acoustic feature of vowel height has long been the object of close scrutiny, research on the vertical representation of vowels in the phonology of English remains inconclusive. By drawing on Sapir's work on phonetic symbolism, this paper empirically investigates the intuitive understanding that SSBE speakers have of their own vowel space. It argues that their implicit knowledge of vowel height differences is best accounted for in terms of a three-tiered vertical axis.

Key words: phonetic symbolism, phonological perception, vowel height.

Phonological framework

Beneath the profound differences separating Jakobson, Fant and Halle (1952) and Chomsky and Halle (1968) lies a shared belief that, logically, distinctive features must be binary. In the case of RP English, the former make no difference between long monophthongs and diphthongs, both types of vowels being assigned a geminate nuclear structure consisting of a) an initial stressed vowel and b) a final unstressed non-syllabic off-glide. Consequently, the phonologically relevant feature in both phonetic oppositions [ɪ]~[i:] and [ʊ]~[u:] is said to be one of nuclear structure (/ 'i/~/ 'iᵿ/, / 'u/~/ 'uᵿ/), not height. As argued by Mendousse (2007), and in keeping with Jakobson's (1949) pledge to Ockham's Razor, the feature [±stressed/unstressed] allowed Jakobson, Fant and Halle (1952) to cut back on 1 of 3 vowel primes and downsize Jones' (1918) set of 20 vowel phonemes to 12.

Chomsky and Halle (1968) in turn rejected Jakobson's creed of parsimony and unwavering belief in the primacy of acoustic cues, which failed to account for the underlying symmetry between palatalisation and front vowels, and velarisation and back vowels. Instead, they provided an articulatory-based variant system where their motor features [±high/low], [±front/back] and [±round/nonround] only partly overlapped with acoustic features [±diffuse/compact], [±grave/acute] and [±flat/natural]. They also postulated the feature [±tense/lax] by virtue of which only underlying tense vowels are made to undergo their advocated Vowel Shift Rule and Diphthongization Rule: surface [ɪ] and [i:] would thus derive from

underlying /i/ and /ē/, in contrast to Jakobson, Fant and Halle (1952) who assigned equal phonological height to both vowels via the phonemic opposition /'i/~/'i̇/; conversely, where the latter posited a phonological height difference in [ε]~[i:] via the phonemic opposition /'ε/~/'i̇/, Chomsky and Halle (1968) placed them on a vertical par via underlying /ε/~/'ε/.

Such differences notwithstanding, both paradigms endorsed a given set of features whose scalar values on the phonetic continuum are filtered down, phonologically, to raw binary oppositions. But the binary doctrine itself is not without its dissenters. In the wake of Ladefoged (1971), who included the feature [±mid] in order to account for a fourth degree of aperture, Singh (1976) assigned /i:, u:/, /ɪ, ʊ/, /ε, ɔ:/, /εɪ, ɒ/ and /æ, ɑ:/ the multi-valued tongue height features [5 high], [4 high], [3 high], [2 high] and [1 high]. Actual tongue height measurements, however, have often been found at variance with their conventional representation on the vowel quadrilateral (Durand 1990). While acknowledging the relevance of tongue height in the study of vowel production, Ladefoged (1971) questioned their value in phonology and recommended primacy be given to the auditory-acoustic correlates of distinctive features given that formant charts yield a more accurate picture of reality than articulatory descriptions. Since then, advocates of the motor theory of speech perception have argued for a more abstract level of articulation, where motor gestures used to decode speech to key articulators correspond to speakers' intentions (Liberman and Mattingly 1985). As noted by Durand (1990: 69), critics have dubbed this readjustment as "the retreat up the vocal tract", but proponents of auditory-acoustic theories in turn are faced with the difficulty that no single cue corresponds to one phonetic category and that acoustic records mesh only indirectly with auditory percepts.

Experimental framework

Heeding the advice of Trubetzkoy (1931), for whom the "linguistic unconscious" should serve as a phonological compass, we undertook to integrate Sapir's (1929) pioneering work on phonetic symbolism within the paradigm of distinctive feature theory. Results of his study have firmly established the unconsciously synaesthetic correlation between speakers' perception of vowel-size differences and actual motor-acoustic (dis)similarities, which Fónagy (1980) later developed into a model of metaphoric ideation where phonetic metaphors are said to be the symbolic projection of speakers' intersensory image of sounds. In the simplest of cases, the actual size or volume of the stimulus is mimetically connected to the articulatory gesture itself, such as the degree of aperture or volume of the resonance chambers.

The scope and cross-cultural consistency of the metalanguage used in phonetics also points to more complex perceptual patterns arising from associative learning mechanisms that correlate size with pitch (Peterfalvi 1970). By allowing for both symbolic responses a) “larger/smaller” vowel ↔ “wider/narrower” ↔ [+low]/[+high] and b) “larger/smaller” vowel ↔ “heavier/lighter” ↔ [+compact]/[+diffuse], phonetic metaphors circumvent the debate on the motor-acoustic content of distinctive features. Furthermore, Taylor and Taylor’s (1962) empirical study demonstrates that speakers are sound-symbolically insensitive to allophonic variations, in which case their vowel-size rankings may be construed as the symbolic projection of their implicit feature knowledge of vowel height differences. In order to assess their mental representation of such differences, participants were thus handed out a vowel-size survey questionnaire and asked, in a forced-choice selection task, to choose from each of 30 nonword minimal pairs the term most likely to mean “smaller” or “larger”.

Results

114 completed questionnaires were collected, yielding the following statistically significant front and back vowel-size rankings: {[i:] < [ɛ, ɛɪ] < [æ, aɪ]}, {[u:] < [ou] < [oɪ, ɔ:]}. These findings strongly support the three-tiered definition of vowel height advocated in binary feature theories, and hence call into question the need, in the name of phonetic realism, for all vowels to be phonologically distinct with respect to height. Back vowel-size rankings, however, remain at odds with Lilly and Viel’s (1998) equal mid-vowel height plotting of {/ɔ:, ɔɪ/}.

Moreover, the non-rejection of the null (random) hypothesis H0 in {[ɪ] ~ [i:], [ʊ] ~ [u:]}, and simultaneous rejection in {[ɪ] ~ [aɪ], [ɛ] ~ [i:], [æ] ~ [ɛɪ], [ɔ] ~ [ou], [ʊ] ~ [ɔ:]}, go against the alleged psychological reality of Chomsky and Halle’s (1968) Vowel Shift Rule: deriving the former from underlying {/i/~ē/, /u/~ō/} incorrectly predicts H0 to be untrue, while deriving the latter from underlying {/i/~ī/, /ɛ/~ē/, /æ/~ǣ/, /ɔ/~ō̄/, /u/~ū/} incorrectly predicts the opposite.

The above results invalidate motor-acoustic height as a phonologically relevant feature in {[ɪ] ~ [i:], [ʊ] ~ [u:]}. Halle and Stevens’ (1969) [±ATR/RTR] feature opposition, which they appended to Chomsky and Halle’s (1968) inventory as an alternative to vowel tenseness in order to account phonologically for a fourth degree of aperture, is an equally unlikely candidate: participants did not respond symbolically to the ATR-associated widening of the pharyngeal cavity in [i:] and [u:] by rating {[ɪ], [ʊ]} and {[i:], [u:]} vowels as being respectively “smaller” and “larger”. Phonemically, such vowel pairs have been alternatively defined as /V/~V:/ (Jones 1918) and /'V/~'V₁/ (Jakobson, Fant and Halle 1952). The

psychological (in)adequacy of each the aforementioned phonological definitions remains a matter for further empirical study.

References

- Chomsky, N., Halle, M. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Durand, J. 1990. *Generative and Non-Linear Phonology*. London: Longman.
- Fónagy, I. 1980. La métaphore en phonétique. *Studia Phonetica* 16, xi-220.
- Halle, M., Stevens, K.N. 1969. On the feature "advanced tongue root". MIT Quarterly Progress Report 94, 209-215.
- Jakobson, R. 1949. On the identification of phonemic entities. In Jakobson, R. 1962, *Selected Writings I*, 418-425. The Hague, Mouton.
- Jakobson, R., Fant, G., Halle, M. 1952. Preliminaries to Speech Analysis: The Distinctive Features and their Correlates. In Jakobson, R. 1987, *Selected Writings VIII*, 585-654. Berlin, Mouton.
- Jones, D. 1918. *An Outline of English Phonetics*. Cambridge: Heffer & Sons.
- Ladefoged, P. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Liberman, A.M., Mattingly, I.G. 1985. The motor theory of speech perception revised. *Cognition* 21, 1-36.
- Lilly, R., Viel, M. 1998. *Initiation Raisonnée à la Phonétique de l'Anglais*. Paris: Hachette.
- Mendousse, K. 2007. Le dilemme de Roman Jakobson face à l'opposition de tension/laxité vocalique. *Histoire Epistémologie Langage* 29.1, 29-68.
- Peterfalvi, J.-M. 1970. *Recherches Expérimentales sur le Symbolisme Phonétique*. Paris: CNRS.
- Sapir, E. 1929. A study in phonetic symbolism. In Mandelbaum, D. G. (ed.) 1949, *Selected Writings of Edward Sapir in Language, Culture and Personality*, 61-72. Berkeley, University of California Press.
- Singh, S. 1976. *Distinctive Features: Theory and Validation*. Baltimore: University Park Press.
- Taylor, I.K., Taylor, M.M. 1962. Phonetic symbolism in four unrelated languages. *Canadian Journal of Psychology* 16, 344-356.
- Trubetzkoy, N.S. 1931. *Die phonologischen Systeme*. *Travaux du Cercle Linguistique de Prague* 4, 96-116.

Prosodic patterns in child speech

Roksolana Mykhaylyk

Center for Advanced Study in Theoretical Linguistics (CASTL), University of Tromsø, Norway

Abstract

This paper presents an experimental study contributing to the issue of prosody-syntax-semantics interaction in child speech. The data from 12 3-4-year-old children acquiring Ukrainian have been analyzed in order to establish main prosodic patterns associated with different semantic types of SVO and SOV sentences. The results show that the children are mostly adult-like in the prosodic realization of post-verbal object pronouns, but often apply non-adult-like prosodic contours to the sentences with contextually-dependent object NPs.

Key words: pitch, prosodic contour, scrambling, Ukrainian

Introduction

It has been established that there is an obvious interaction between prosody, semantics and syntactic structure in adult speech (Vallduvi 1992, Cinque 1993, Selkirk 1995, Zubizarretta 1998, *inter alia*). In child language, however, the role of prosody has not been clearly defined. Some studies suggest that young children differ from adults in the use of intonation/pitch: i.e., the child variations in the prosodic contours are often due to different emotional and interactional contexts, and not to the semantics or information structure (Behrens & Gut 2005, Chen & Fikkert 2007). Other studies (e.g., Chen 2010 & 2011) show that 4-5-year-olds are able to use intonation to encode informational status of arguments. A specific issue to clarify concerns a prosody-syntax-semantics interaction in intonational languages with a ‘flexible’ word order since in such languages, scrambling might have the same interpretational effect as prosodic destressing (Neeleman & Reinhart (1998)).

Ukrainian: Syntax-prosody Interaction

Prosodic and syntactic properties of Ukrainian interact in an interesting way. The unmarked prosodic structure of Ukrainian is shown in (1) (as defined by Féry et al. (2007)):

- | | | | |
|-----|----------------------------------|------------|-----------|
| | LH* | | HL* |
| (1) | [[DIVčyna]p | [čytaje | roMAN]p]i |
| | girl | is-reading | novel |
| | ‘A/the girl is reading a novel.’ | | |

The basic syntactic structure of Ukrainian is SVO. However, pronouns (and other context-dependent elements, such as definite or partitive NPs) usually

must appear in pre-verbal position ((2) & (4)). This requirement can be circumvented only if the intonation of a sentence is modified (as in (3) & (5)).

- | | | | | | | |
|--------------------------------|--------------|------------|-----|-------------|-----------|--------|
| (2) Divčyna | <u>roman</u> | pročytala. | 000 | (3) Divčyna | PROČYTALA | roman. |
| girl | novel | read | | girl | read | novel |
| 'The girl has read the novel.' | | | | | | |
| | | | | | | |
| (4) Divčyna | <u>joho</u> | pročytala. | | (5) Divčyna | PROČYTALA | joho. |
| girl | it | read | | girl | read | it |
| 'The girl has read it.' | | | | | | |

This research investigates whether, in cases where children fail to move a contextually-dependent noun or pronoun, they will compensate for this lack of movement by destressing the object and applying the falling pitch accent on the verb, as is shown for adults in Antonyuk-Yudina & Mykhaylyk (2009). If children are able to establish context relatedness for the direct object, they will apply a distinct prosodic contour to the SVO sentences with the definite/partitive/pronominal objects as compared to the neutral contour associated with the indefinite object in the SVO sentence.

Experiment

Method

Twelve 3-4-year-old monolingual Ukrainian children participated in the experiment. The method was a picture description task, in which the pictures and questions were designed to elicit particular types of nouns/pronouns. The design was based on Mykhaylyk (2010), and the conditions matched those presented in Antonyuk-Yudina & Mykhaylyk's (2009) study on adult Ukrainian. The goal was to identify the intonation preferred for two types of syntactic structures (i.e., SVO and SOV) in four contexts triggering use of definite, partitive, indefinite or pronominal direct object. For example, in the Definite/Pronominal Context, there was a car in one picture and Kangaroo with the same car and a set of tools in another picture, and the expected responses were either 'Kangaroo her/the car is fixing' or 'Kangaroo IS FIXING her/the car'. Target sentences were analyzed acoustically and labeled using ToBI labeling conventions (Pierrehumbert 1980, Silverman et al. 1992). The key principles of the child data analysis were identical to the analysis of adult results: the common types of prosodic contours were identified, and the group results for the object stress and the verb pitch type were analyzed.

Results

The results show that the children distinguish several types of prosodic contours and are able to use them in appropriate contexts. The basic SVO structure with an indefinite object has an unmarked prosodic contour (as in (1)). On the other hand, the SVO structure with a pronominal direct object has the strongest falling pitch accent realized on the verb (H^*+L), while the object is prosodically destressed (see Fig.1), exactly as in the adult results (see Antonyuk-Yudina & Mykhaylyk 2009).

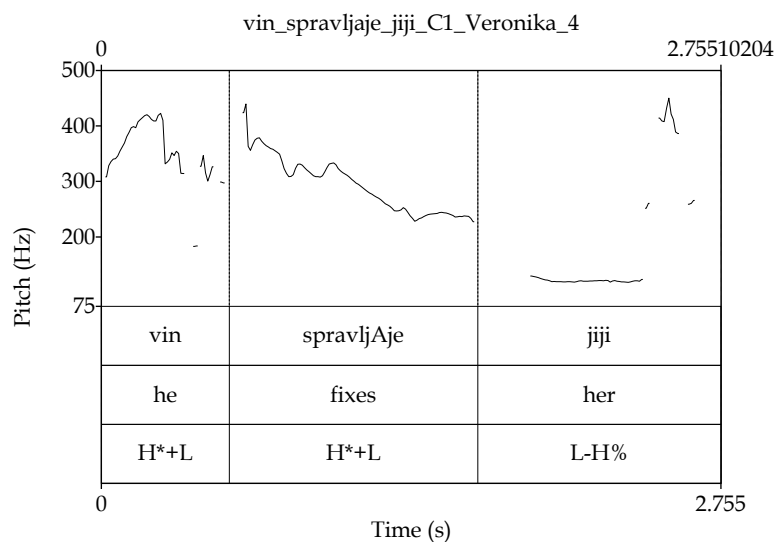


Figure 1. Prosodic contour of an SVO structure with a pronominal object

The children exhibited more variability in SVO sentences with contextually-dependent full NPs. When they dropped a subject and used a two-word structure with a definite object in a postverbal position, their prosody was mostly adult-like: with a falling pitch accent on the verb and a destressed object. However, the three-word structures often had unmarked prosody regardless of the context.

The total group results confirm that the children behave mostly adult-like in never stressing indefinite direct objects in the basic SVO structure (cf. Table 1 and Table 2). They also closely approximate adults by destressing pronouns at 75% (cf. 100% for adults in Table 2), but are less adult-like in the prosodic realization of sentences with other context-dependent direct objects. As shown in Table 1, the child prosody of sentences with definite and partitive objects differs considerably from the adult prosody in similar semantic contexts (Table 2).

Table 1. SVO structures with four types of direct objects: Children, %.

	Pronominal object	Definite object	Partitive object	Indefinite object
Destressed object	75	50	75	0
Falling pitch on verb	100	56	50	50

Table 2. SVO structures with four types of direct objects: Adults (From Antonyuk-Yudina & Mykhaylyk (2009)), %

	Pronominal object	Definite object	Partitive object	Indefinite object
Destressed object	100	87	75	0
Falling pitch on verb	100	100	88	25

While the child data exhibit a clear contrast between the stress placement on the direct object (definite and partitive objects are destressed more often (50% & 25%) than indefinite objects (0%)), there is no such contrast for the prosodic realization of the verb. The definite and partitive contexts do not trigger use of the falling pitch on the verb more often than indefinite contexts (56%, 50% and 50%, respectively). It is important to mention, however, that none of the children's prosodic contours could be considered 'incorrect', but some of them just differed from those normally used by adults.

Conclusion

These findings are comparable to Chen (2011): both studies show that children at the age of 4 are already sensitive to the context-prosody correlation. However, the children's prosodic patterns appear to be more variable than the adult patterns in similar semantic contexts. While in this study the child and adult data could not be compared directly due to the difference in methodology, established patterns deserve further investigation.

References

- Antonyuk-Yudina, S., R. Mykhaylyk. 2009. Prosodic effects in word order: Evidence from Ukrainian. Presentation at NELS-40, MIT.
- Chen, A. 2011. Tuning information packaging: Intonational realization of topic and focus in child Dutch. *Journal of Child Language*. Available on CJO 2006.
- Chen, A., Fikkert P. 2007. Intonation of early two-word utterances in Dutch. In Trouvain, J., William, J.B. (eds.), *Proc. of 16th ICPHS*. Pirrot GmbH, Dudweiler.
- Féry, C., Paslawska, A., Fanselow, G. 2007. Nominal split constructions in Ukrainian. *Journal of Slavic Linguistics* 15 (1), 3-48.
- Mykhaylyk, R. 2010. Optional object scrambling in child and adult Ukrainian. PhD dissertation, Stony Brook University.
- Neeleman, A., Reinhart, T. 1998. Scrambling and the PF interface. In Butt, M., Geuder, W. (eds.), *The projection of arguments: Lexical and compositional factors*, 309-353. Stanford, CA: CSLI Publications.
- Zubizarreta, M.L. 1998. *Focus, prosody, and word order*. Camb., MA: MIT Press.

Tonal and syntactic correlates of focus perception in Greek and Russian

Olga Nikolaenkova

Lab of Phonetics & Computational Linguistics, University of Athens, Greece

Abstract

The present paper reports on the way word order and tonal slope influence focus perception in Russian and Greek declaratives. Perception experiment was carried out in order to identify the impact falling tone slope and word order change have on focus identification. The speech material included manipulated stimuli with 5 and 3 different falling tone slopes for Russian and Greek respectively and involved 6 different word orders. The results of the perception test indicate that: (1) slope manipulation is a reliable focus indicator; (2) word order change is not sufficient for focus identification; (3) no major cross-language contrasts were observed.

Key words: focus, perception, tonal slope, Russian, Greek

Introduction

Local tonal expansion accompanied by post-focal tonal compression is widely accepted as one of the main tonal correlates of focus (Botinis 2003, Nikolaenkova 2010, Xu 2011). Having a closer look at focus production data though we identify focus-marking through gradient means such as duration, F0 timing etc.

The goal of the present study is to reveal the complex nature of focus implementation strategy. Contrastive analysis of the tonal slope and word order change effects on focus perception in Russian and Greek will help us estimate how deep the influence of linear hierarchy on focus perception is.

Methodology

The experiment was based on an exclusively designed computer application. Each listener was instructed to click on the most prominent unit if any. The time interval between the stimuli was 1,5 sec. The experiment consisted of five sets, 18 stimuli each, for Russian and four sets for Greek repeated 10 times in random order. For Russian the experiment yielded 9000 responses (90 utterances x 10 repetitions x 10 informants). For Greek the experiment yielded 7200 responses (72 utterances x 10 repetitions x 10 informants).

Speech material

Speech material for the experiment consisted of a Russian declarative sentence «*ma-RI-na MJA-la man-da-RI-ny* / Marina was pressing the mandarins» and a Greek one «*i me-LI-na MA-lo-ne ti-MA-na-mu* / Melina was arguing with my mother» (block letters mark the stressed syllables). Female native speakers of Russian and Greek respectively were recorded in

a sound treated booth at the University of Athens Phonetics Laboratory. The task was first to read the phrase aloud and then to answer three questions eliciting three focus placements S-focus, VP-focus and O-focus always using exactly the same phrase. The speech material was recorded directly to computer disk using the PRAAT software package.

Focus production data

According to the production data from Russian and Greek (see Table 1) we should highlight the following:

focus was found to significantly influence both rising and falling tone speed in Greek while in Russian its impact was observed only in case of S-focused utterance;

no direct linkage between focus application and rising-falling tone speed balance has been revealed;

VP-focus has been found having the opposite effect on tonal speed range as well as on the balance between rising and falling speed in both languages.

Table 1. Tonal rise and fall speed rates within the stressed syllable and the one following it as a function of focus placement (+/- focus) in Russian and Greek.

Hz/s	Subject				Verbal Phrase				Object			
	N		F		N		F		N		F	
	Ru	Gr	Ru	Gr	Ru	Gr	Ru	Gr	Ru	Gr	Ru	Gr
Rise	682	51	668	643	286	62	286	349	330	152	330	165
Fall	125	307	668	1240	582	242	582	173	422	242	422	591

Manipulation

In accordance with the goals of the perception experiment naturally produced neutral utterances went through tonal slope manipulations and word order manipulations.

Tonal slope has been edited within the first post-stressed syllable through adjusting the speed of the tone. This set consisted of fifteen and nine stimuli for each utterance in Russian and Greek respectively. The slope observed for the falling tone in cases of natural focus productions was the reference point for further 30% step increase and decrease aimed at perception threshold identification. Word order manipulated stimuli involved lexical units being

cut and dislocated. All the manipulated material went through tonal normalization.

Results and Discussion

The perception test showed that in total focus identification was 100% in Russian while in Greek it was 33% (see Table 2). The reason for such contrast could be the fact that the Russian part involved two more slopes.

It is remarkable that in both languages successful identification rates were higher for V focus, while S focus had the lowest identification rate (see Table 2). This finding disagrees with our previous focus perception tests where VP and O focus (Nikolaenkova 2009). Having performed one-way ANOVA test with focus type as independent variable, we found that the effect of focus type on successful identification was statistically insignificant for both languages ($p > 0.05$).

Table 2. Total successful focus identifications as a function of focus type in Russian and Greek.

Focus Type	S		V		O	
	Rus	Gr	Rus	Gr	Rus	Gr
Success	2110	400	4150	1040	2740	940
% Success	23.44	5.56	46.11	14.44	30.44	13.06

On the other hand distribution of the parameters used in manipulating our stimuli neither revealed any tendency (see Figure 1). Nevertheless slope parameter, if accounted in total, has a significant domination of 74.84% ($p < 0.01$) and 15% ($p < 0.01$) in Russian and Greek respectively.

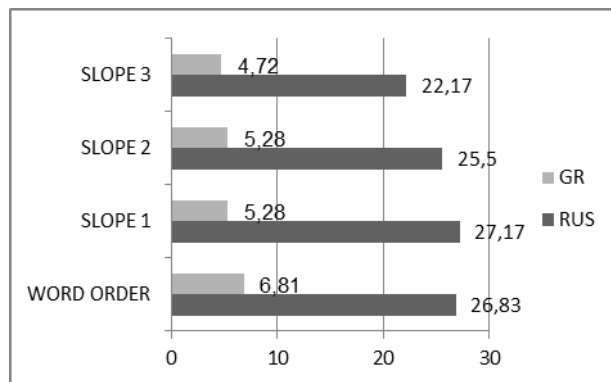


Figure 1. Successful focus identification percentages as a function of manipulation parameters in Russian and Greek.

Having analysed the impact of word order and slope on focus identification separately we should acknowledge the leading role of slope in sounding speech. This conclusion is supported not only by the difference in successful identification rates of word order and slope taken separately but also by almost no change in identification rates in cases when both correlates have been applied. This almost invisible position of word order manipulation performs a strong argument in favour of ending the prevailing linear perception in studying focus.

Acknowledgements

I would like to thank P. Karamolenkos for the programming and testing of the relative application. My deep gratitude to all those who participated in the experiment and to those who gave me their kind guidance and advisory.

References

- Botinis, A. 2003. Focus and gender tonal interactions. XVth International Congress of Phonetic Sciences, 1826-1829. Barcelona, Spain.
- Nikolaenkova, O. 2009. Rhythmic and tonal correlates of focus in Greek. Master dissertation. University of Skovde, Sweden.
- Nikolaenkova, O. 2010. Perception of «tonal focus» in Greek. Botinis, A. (ed.), Proc. of 3rd ISCA Workshop on Experimental Linguistics ExLing 2010, 121-124.
- Xu, Y., Chen, S., Wang, B. Forthcomming. Prosodic focus with and without post-focus compression: A typological divide within the same language family?

Production of Greek and Turkish vowels by bilingual speakers

Elina Nirgianaki¹, Ougour Chasan¹, Evgenia Magoula²

¹Laboratoy of Phonetics & Computational Linguistics, University of Athens, Greece

²Department of Education, University of Athens, Greece

Abstract

The present study examines the acoustic vowel space of Greek and Turkish vowels produced by bilingual speakers. The results are presented for both languages' vowels production by bilingual speakers, as well as compared with the vowels production by monolingual speakers. It is revealed that the acoustic space of the Greek vowels produced by bilingual speakers is larger than the acoustic space of their Turkish vowels, mainly due to the lower and more front position of the point vowels [a] and [i], respectively. Moreover, the acoustic space of the Greek vowels produced by bilingual speakers is smaller than the one of the Greek vowels produced by monolingual speakers. On the contrary, the acoustic space of the Turkish vowels produced by bilingual speakers is larger than the one of the Turkish vowels produced by monolinguals.

Key words: Greek, Turkish, bilinguals, vowels, formants

Introduction

A number of acoustic studies on Greek vowels have been carried out in the past (i.e. Jongman et al, 1989; Hawks et al, 1995; Fourakis et al, 1999) and there are also a few studies examining the acoustics of Turkish vowels (e.g. Türk et al, 2004; Kiliç, 2003). However, there have been no published results on vowels produced by bilingual speakers of Greek and Turkish. The present study examines the acoustic vowel space of Greek and Turkish vowels produced by bilingual speakers, and compares these findings with respective results from monolingual Greek and Turkish speakers.

Methodology

Four speakers, two females and two males, recorded the experimental material. All are bilingual speakers of standard Greek and Turkish, who live in Greece, and none of them has any history of speech or hearing disorders. The five Greek vowels [a], [e], [o], [i], and [u] and the eight Turkish vowels [a], [e], [ɯ], [i], [o], [œ], [u], and [y] were recorded in real, two-syllable words (CVCV). Each vowel was in initial, stressed syllable of the word and the previous and following consonants were stops (except for the Turkish [a] and [y], which were preceded by [m] and [f] and followed by [s] and [z], respectively). The carrier phrase for Greek was 'i 'leksɪ _ 'ine elini'ci' (The word _ is Greek) and for Turkish "bir ta'ne ... al'mak i'stijorum" (I want one _ to buy). Each token was repeated once, yielding a total of 52 tokens.

For the comparison between the Greek vowels' production by bilingual and monolingual speakers, four Greek monolingual speakers, two females and two males, recorded the same Greek experimental material. For the comparison between the Turkish vowels' production by bilingual and monolingual speakers, the data from Türk et al (2004) has been used (the averaged *F1* and *F2* values of the Turkish vowels produced by 15 male and 14 female speakers in words and sentences).

F1 and *F2* were measured by hand marking from wideband spectrogram, in the middle of each vowel. The statistical analysis was carried out in Excel.

Results

Table 1 and table 2 present the *F1* and *F2* values for the Greek and Turkish vowels, respectively, produced by the bilingual speakers.

Table 1. Mean *F1* and *F2* values (in Hz) for Greek vowels, as a function of speaker gender.

Vowel	Male		Female		Mean	
	<i>F1</i>	<i>F2</i>	<i>F1</i>	<i>F2</i>	<i>F1</i>	<i>F2</i>
[a]	809	1387	836	1333	822	1360
[e]	574	1713	619	2210	596	1962
[i]	411	2002	474	2590	443	2296
[o]	520	1044	610	1180	565	1112
[u]	411	981	574	1125	492	1053

Table 2. Mean *F1* and *F2* values (in Hz) for Turkish vowels, as a function of speaker gender.

Vowel	Male		Female		Mean	
	<i>F1</i>	<i>F2</i>	<i>F1</i>	<i>F2</i>	<i>F1</i>	<i>F2</i>
[a]	673	1234	701	1469	687	1351
[e]	493	1777	545	2201	519	1989
[ɯ]	483	1369	492	1731	488	1550
[i]	420	1849	465	2219	443	2034
[o]	502	836	664	1035	583	935
[œ]	538	1397	538	1821	538	1609
[u]	438	990	456	1179	447	1085
[y]	411	1623	447	1948	429	1785

Regarding the formant values of the five vowels that are common in Greek and Turkish, it is worth to mention that the Greek [a] has around 20% higher *F1* than the Turkish [a] and the Greek [i] and [o] have around 11.5% and 16%, respectively, higher *F2* than the Turkish ones.

Figures 1 and 2 show $F1 \times F2$ acoustic spaces for Greek and Turkish vowels, respectively. In each space, the vowels are plotted by the mean frequencies of their formants. It is illustrated that the Greek [a] is lower than the Turkish [a] and the Greek [i] more front than the Turkish one. Both Greek and Turkish [e] and [u] lie almost in the same space, and the three Turkish vowels [ɯ], [œ] and [y] lie in the mid vowel space, more close to the front than the back vowels. The lower and more front position of the point vowels [a] and [i], respectively, define a larger acoustic space of the Greek vowels than the acoustic space of the Turkish vowels.

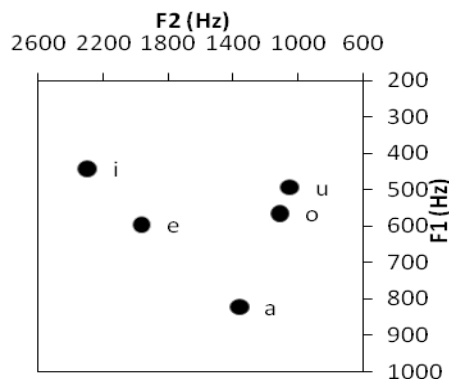


Figure 1. The Greek vowels plotted in an $F1 \times F2$ space (mean values).

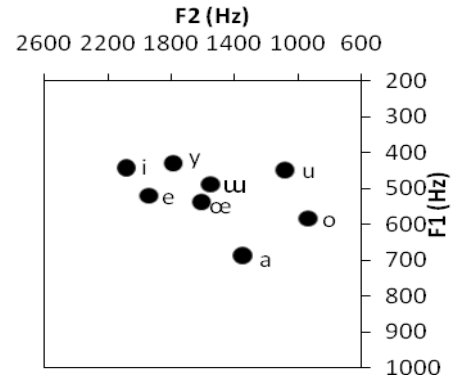


Figure 2. The Turkish vowels plotted in an $F1 \times F2$ space (mean values).

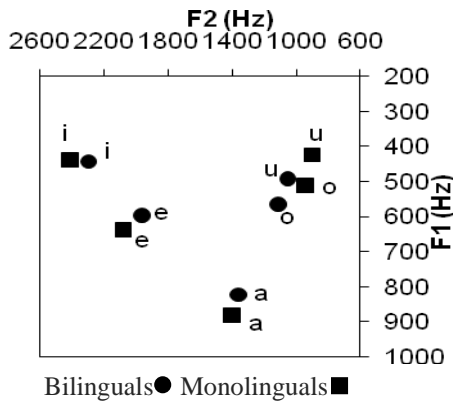


Figure 3. The Greek vowels produced by monolingual and bilingual speakers plotted in an $F1 \times F2$ space (mean values).

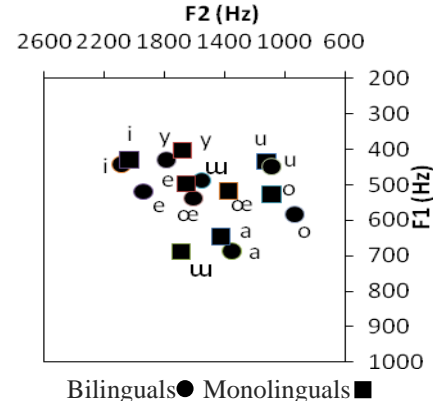


Figure 4. The Turkish vowels produced by monolingual and bilingual speakers plotted in an $F1 \times F2$ space (mean values).

In Figures 3 and 4 the *F1* and *F2* frequencies of the Greek and Turkish vowels produced by bilingual speakers are plotted along with the results from monolingual Greek and Turkish speakers, respectively.

Regarding the Greek vowels, it is clearly shown that all vowels produced by monolingual speakers occupy more extreme positions in the acoustic space than the ones produced by bilingual speakers. On the other hand, the Turkish vowels produced by monolingual speakers lie either in about the same ([i], [u], [y]) or in more central position ([a], [o], [e], [œ]) than the ones produced by bilinguals.

Conclusions

The results of the present study provide important cues for the production of Greek and Turkish vowels by bilingual speakers.

First, the acoustic space of the Greek vowels produced by bilingual speakers is larger than the acoustic space of their Turkish vowels. This result accords with the results of Jongman et al. (1989), Hawks et al. (1995) and Fourakis et al (1999) in that the Greek vowels are well separated in an acoustic space, allowing for maximal contrast between vowel categories. In addition, the acoustic space of the Greek vowels produced by bilingual speakers is smaller than the one of the Greek vowels produced by monolingual speakers, whereas the acoustic space of the Turkish vowels produced by bilingual speakers is larger than the one of the Turkish vowels produced by monolinguals. This provides the preliminary evidence that due to the production of both vowel systems, the bilingual speakers tend to equalize the one with the other in terms of their acoustic space.

References

- Fourakis M., Botinis, A., Katsaiti, M. 1999. Acoustic characteristics of Greek vowels. *Phonetica* 56, 28-43.
- Hawks, J.W., Fourakis M.S. 1995. The perceptual vowel spaces of American English and modern Greek: a comparison. *Language and Speech* 38, 237-252.
- Jongman, A., Fourakis, M., Sereno, J. 1989. The acoustic vowel space of Modern Greek and German. *Language and Speech* 32, 221-248.
- Kiliç, M.A. 2003. Türkiye Türkçesi'ndeki ünlülerin sesbilgisel özellikleri. In: Özsoy, A.S., Taylan, E.E., Aksu-Koç, A., Akar, D., Nakipoğlu, M. (Eds.), *Studies in Turkish Linguistics*. Boğaziçi University Press, Istanbul, pp. 3-18.
- Türk, O., Şayli, Ö., Özsoy, A. S., Arslan, L. M. 2004. Türkçede Ünlülerin Formant Analizi. *Proceedings of the 18th National Conference in Turkish Linguistics*. Ankara University.

The validity of some acoustic measures to predict voice quality settings: trends between acoustic and perceptual correlates of voice quality

Luiz Carlos Rusilo¹, Zuleica Antonia de Camargo², Sandra Madureira²

¹ Department of Actuarial and Quantitative Methods-DAMQ-Pontifical Catholic University of São Paulo, Brazil

² Integrated Acoustic Analysis and Cognition Laboratory-LIAAC- Pontifical Catholic University of São Paulo, Brazil

Abstract

The present study aims at investigating the power of acoustic measures to predict voice quality settings detected perceptually. The corpus was composed by speech samples recorded by 60 subjects. The audio samples were analyzed in PRAAT by means of the SG Expression Evaluator Script (Barbosa, 2009), which extracts f0 and f0 first derivate, intensity, spectral tilt and long-term average spectrum measures. The same samples were perceptually evaluated by means of the Voice Profile Analysis Scheme for Brazilian Portuguese (BP-VPAS). The methodological procedures made it possible to identify the clusters related to the acoustic measures group and that of the perceptual judgement results (cluster analysis) as well as to correlate these two groups (canonic correlation analysis).

Key words: speech acoustics, voice quality, perception, statistical analysis

Introduction

Many conditions that influence voice quality analysis are not addressed in studies that search for acoustic-perceptual correlations (Hammarberg, Gauffin, 1995). Furthermore, there are few studies describing the acoustic correlates of voice quality settings which can be identified perceptually with the help of the Voice Profile Analysis Scheme (VPAS) (Laver, Mackenzie-Beck, 2007).

This study aims at investigating the power of some acoustic measures (fundamental frequency-f0 and f0 first derivate, intensity, spectral tilt and long-term average spectrum-LTAS) to predict voice quality settings detected perceptually, by means of cluster analysis and canonic correlation analysis. The detailed description of acoustic correlates and its discriminatory power to predict the perceived voice quality settings is not an easy task, but it can provide a detailed description of events related to voice quality settings in laryngeal, supralaryngeal and tension domains.

Methods

The corpus was composed by semi-spontaneous speech samples and repetitions of 10 key-sentences samples (Camargo, Madureira, 2008), recorded in a radio studio by 60 subjects. The audio samples were analyzed by means of the SG Expression Evaluator, a revised version of the script developed by Barbosa (2009), running in the software PRAAT, which automatically extracts acoustic measures: f0 (median, inter-quartile semi-amplitude, skewness and 0,995 quantile) and its first derivate (df0- mean, standard-deviation and skewness), intensity (skewness), spectral tilt (SpTt – mean, standard-deviation and skewness) and long-term average spectrum (LTAS frequency standard-deviation).

Some of the measures were normalized during the application of the script, such as f0 and df0, SpTt and LTAS frequency. The same samples were perceptually evaluated by two expert subjects, using the Voice Profile Analysis Scheme for Brazilian Portuguese: BP-VPAS (Camargo, Madureira, 2008, 2010).

The acoustic measures and voice quality settings judgments were statistically analyzed by means of cluster analysis and canonic correlation analysis (Lattin et al, 2011) in order to investigate the validity of each acoustic parameter to predict the voice quality settings which had also been evaluated perceptually. The cluster analysis was applied intragroup and the canonic analysis intergroups. There were two groups under analysis: the acoustic measures and perceptual judgement results. To derive the statistical measures the *software* Xlstat from Addinsoft was used.

Results

The results point to some correlations between acoustic measures concerning f0, df0, intensity, SpTt and LTAS and the perceived voice quality settings.

The cluster analysis applied to perceptual data yielded four classes: Class 1 (70,1%) - modal (SD 1,3%, CI 67,6% to 72,7%), Class 2 - laryngeal hyperfunction, and harsh voice (SD 1,1%, CI 15,2% to 19,4%), Class 3 - pharyngeal expansion, lowered larynx and creaky voice (SD 0,8%, CI 7,0 to 10,1%) and Class 4 - spread lips, advanced tongue tip, pharyngeal constriction, vocal tract hyperfunction, creaky voice, whisper and harsh voice (SD 0,6%, CI 2,9 to 5,1%).

The cluster analysis applied to acoustic data yielded six classes, two of them comprising a large number of observations: Class 1 (20,80%) and Class 2 (76,11%). They were discriminated by df0-mean and skewness measures. The other four classes (3,09%) were discriminated by f0 and df0 measures. Figure 1 presents the dendrogram related to the perceptual measures and Figure 2 the correlations between acoustic and perceptual data.

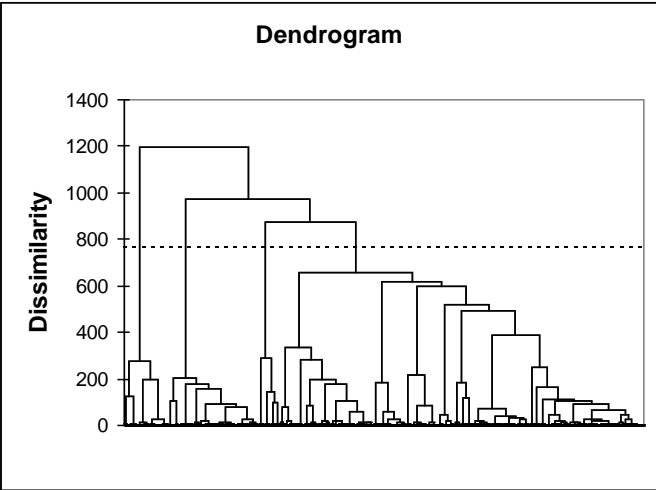


Figure 1. Dendrogram of perceptual data related to voice quality judgements (BP-VPAS).

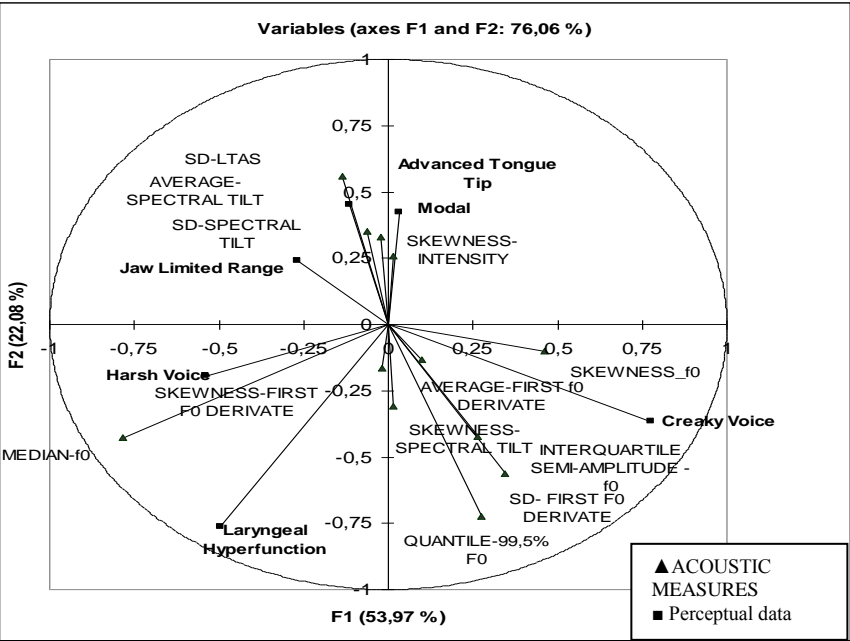


Figure2. Correlations between acoustic and perceptual data.

The correlations shown in Figure 2 concern the most frequent voice quality settings (jaw minimized range, advanced tongue tip, laryngeal hyperfunction, modal, creaky and harsh voice) and acoustic measures.

Conclusions

The methodological procedures made it possible to identify the clusters related to the acoustic measures group and that of the perceptual judgement results (cluster analysis) as well as correlate these two groups (canonic correlation analysis).

Acknowledgements

We acknowledge Plínio Barbosa from UNICAMP for the revised version of the SG Expression Evaluator Script (2009).

References

- Barbosa, P.A. 2009. Detecting changes in speech expressiveness in participants of a radio program In Proc. of Interspeech. v.1, 2155-2158. Brighton, Reino Unido.
- Camargo, Z., Madureira, S. 2008. Voice quality analysis from a phonetic perspective: Voice Profile Analysis Scheme Profile for Brazilian Portuguese (BP-VPAS). In Proc. of the 4th Conf on Speech Prosody v.1, 57–60. Campinas, Brazil.
- Camargo, Z., Madureira, S. 2010. The acoustic analysis of speech samples designed for the Voice Profile Analysis Scheme for Brazilian Portuguese (BP-VPAS): long-term f0 and intensity measures In Proc. of the 3rd ISCA Workshop on Experimental Linguistics ExLing 2010, 33-36. Athens, Greece.
- Hammarberg, B., Gauffin, J. 1995. Perceptual and acoustics characteristics of quality differences in pathological voices as related to physiological aspects. In Fujimura, O., Hirano, M (eds.), Vocal fold physiology, 283-303. San Diego, Singular Publishing.
- Lattin, J., Carrol, Douglas J.D., Green, P.E. *Análise de dados multivariados*. São Paulo: Cengage Learning, 2011.
- Laver, J; Mackenzie-Beck J. 2007. Vocal Profile Analysis Scheme-VPAS. Queen Margareth University College-QMUC, Speech Science Research Centre, Edinburgh.

Lexical stress in Modern Halh Mongolian

Yumei Sang

Department of Linguistics, Laboratory CLILLAC-ARP, Paris-Diderot University,
France

Abstract

The purpose of this paper is to study lexical stress in Mongolian and its place of realization. Words and short declarative sentences are analyzed. We observe certain regularity in the rising of F0 on the most prominent syllables. We also show that duration is used as a correlate of stress in the sense that long vowels are privileged to receive stress: the stress falls on the syllable which contains a long vowel whatever its place within a word. In the case there is no long vowel, the stress falls on the second syllable. (Stress design here lexical stress).

Key words: Speech, prosody, stress, Mongolian, Halh

Introduction

The study focuses on the variety of Mongolian that is spoken in the Republic of Mongolia. There are different opinions on the nature and place of stress in Mongolian. Two prosodic units, the *syllable* and the *more* have been said to carry the stress:

1. Stress falls on the first syllable (Luvsanvandan 1976, 1986, Moomoo 1979).
2. Stress falls on the first syllable if it contains a long vowel. If the first syllable does not contain a long vowel, the stress falls on the second syllable (Harnuad2003).
3. There is no stress in Mongolian on the lexical level. Vowel harmony works as a rhythmic feature. A mora analysis is proposed (Svantesson and al. 2005, Karlsson 2005, Svantesson and Karlsson 2004).
4. We show that the stress falls on the syllable which contains a long vowel. In the case there is no long vowel, the stress falls on the second syllable.

The only consensus is that lexical stress is not distinctive in Mongolian: no minimal pairs can be found. In this paper, we will discuss these theories by studying the stress in Mongolian. The analysis takes into account a test that we have performed concerning the perception of the most prominent syllables. We will also discuss the length of the vowel.

Corpus analysis

The analysis is performed on a corpus composed of more than 70 single words and 50 short declarative sentences read by two Mongolian native-speakers. For the corpus analysis, we used the signal processing software WinPitch.

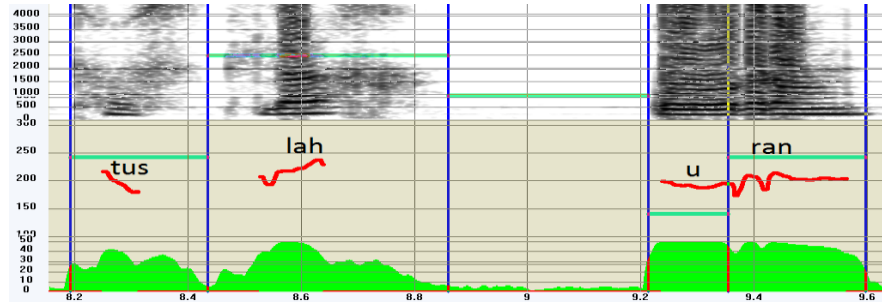


Figure1. Spectrogram of two utterances: “tuslah” (to help) and “uran” (artistic).

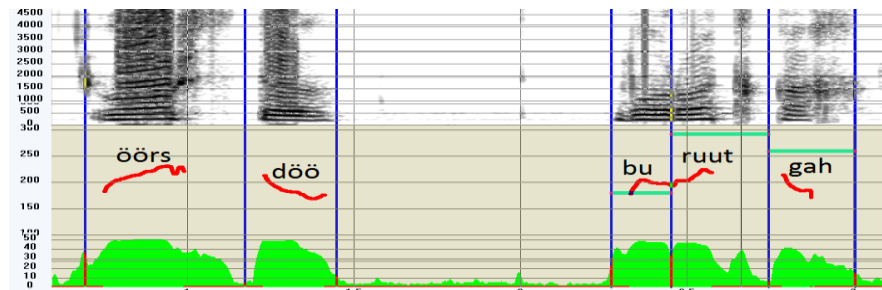


Figure2. Spectrogram of two utterances: “öörsdöö” (self) and “buruutgah” (to accuse).

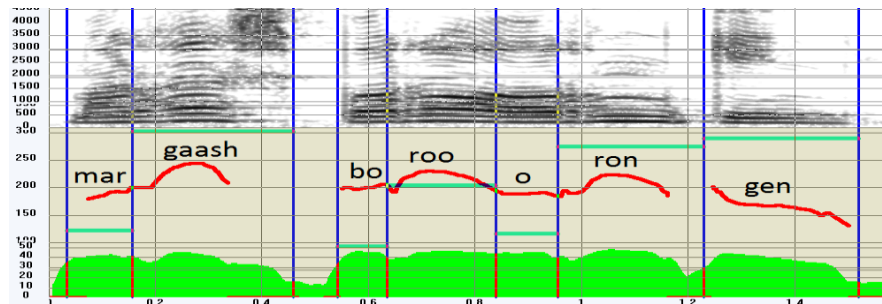


Figure3: Spectrogram of the utterance: “Margaash boroo oron gen” – Tomorrow it will rain. Margaash-tommorw, boroo-rain, oron-fall.FUT. gen-final partical.

Results and discussion

The utterances in the Figure1, do not contain any long vowel. We see that the stress falls on the second syllable both for “*tuslah*” (to help) and “*uran*” (*artistic*). In Figure2, we showed two words each containing a long vowel, “*öördöö*” (self) and “*buruutgah*” (to accuse). We see that there is a F0 rising on the first syllable of “*öördöö*”, which is a long vowel, and the second syllable of “*buruutgah*” which is also a long vowel. In Figure3, “*Margaash boroo oron gen*” (tomorrow it will rain), the spectrogram shows that the stress falls on the second syllable of “*Margaash*” and “*Boroo*” which both contain a long vowel. In “*oron*”, there is no long vowel; we see a F0 rising on the second syllable. So in Figures 1 to 3, we see a F0 rising on the most prominent syllables. The priority place of stress is the syllable containing a long vowel and they may be in any position within a word. We also see that if a word contains no long vowel, the stress falls on the second syllable. This is observed throughout the entire corpus, with few exceptions probably due to word internal structure of these words. A more detailed study would be needed to explain these observations.

We believe that the historical development of vowels in Mongolian makes it difficult to investigate lexical stress, in the sense that Mongolian has undergone a major change at the lexical and phonetic level with arrival of the Cyrillic alphabet. In modern Mongolian, it seems that non-initial vowels, long or short, tend to be shortened. However, initial vowels keep their length because of their strong position in vocalic harmony. These length variations make it more difficult to analyze lexical stress in Mongolian. We think it could be why some previous theories (Karlsson, 2005), (Svantesson, 2005) have assumed that long vowels only appear in the initial position. Karlsson (2005) performed an investigation of vowel length, the results of which showed that a short vowel in the initial syllable has a length 60% that of a long vowel and a short vowel in the non-initial syllable has a length 72% that of a long vowel. These results do not conform to the principle of vocalic harmony which states that the initial vowel has a strong position because of its role of vowel assimilation of following syllables. We believe that the reason why the percentage of initial short vowel is lower than the percentage of the non-initial short vowel, (60% vs 72%), could be that the non-initial short and long vowels were mixed in the analysis. Even if it is sometimes shortened, we think that the long vowel in non-initial syllable stays phonologically long because it receives the stress.

We also believe that the vowel length variation could explain these different theories on the lexical stress in Mongolian.

Conclusion

In this paper, we have showed that the lexical stress is conventional in Mongolian and that its place of realization is governed by certain rules. A syllabic analysis is proposed: lexical stress falls in priority on the syllable containing a long vowel, regardless its position within a word. When there is no long vowel, stress falls on the second syllable. We showed that F0 rising is the most correlated patterns to signal prominence. We discussed also the duration of vowels in initial and non-initial position in the word, especially the status of non-initial reduced long vowels. We think that these vowels are phonologically long, since they receive the stress.

References

- Harnaud, H. 2003. A Basic study of Mongolian prosody. Publications of the Department of Phonetics, University of Helsinki 45. Helsinki.
- Karlsson, A.-M. Svantesson, J.-O. 2004. Prominence and mora in Mongolian. *Speech Prosody*, 65-58.
- Karlsson, A.-M. 2005. Rythme and intonation in Halh Mongolian. Lund University.
- Luvsanvandan, Sh. 1976. Orchin tsagiin mongol helnii butets(2). M&D.
- Luvsanvandan, Sh. 1986. Orchin tsagiin mongol helnii butets(1). M&D.
- Moomoo, S. 1979. Mongol helnii avia züi. MUIS.
- Svantesson, J.-O. Tsendina, A., Karlsson, A., Franzén, V. 2005. The phonology of Mongolian. Oxford University Press.

Pretonic schwa reduction in Dutch: Frequency effects interact with phonotactics

Marjoleine Sloos

Germanic Department, Freiburg University, Germany

Abstract

Word frequency plays a role in reduction processes: high-frequency words undergo more reduction than low-frequency words. Pretonic schwa reduction in Dutch is susceptible to frequency effects in this way, but the well-formedness of the cluster that remains after schwa deletion also plays a role. An experiment shows that cluster well-formedness and frequency effects in fact interact with each other. This suggests that phonological models should incorporate both lexical as well a grammatical information.

Key words: Frequency effects, phonotactics, pretonic schwa, reduction.

Introduction

It has been observed for Dutch, that schwa deletion is more likely to be applied, when the resulting consonant cluster consists of an obstruent and subsequent liquid (Booij 1995). Thus all other things being equal, *gelijk* [χə'leik] 'even' is expected to undergo schwa deletion more often than *geniep* [χə'ni:p] 'mean', because in Dutch χl is a well-formed cluster but χn is not. However, pretonic vowels in Dutch are also known to be more reduced in high-frequency (HF) words than in low-frequency (LF) words (van Oostendorp 1999). This raises the question of whether frequency effects interact with the well-formedness of the cluster that remains after deletion. Is it the case that frequency effects are only found in potential well-formed onset clusters? Or is it rather the case that schwa reduction in consonant-liquid clusters is dependent on word frequency? How does frequency interact with phonotactics exactly?

To investigate the effect of cluster well-formedness (CWF) and frequency, a word reading task was carried out. HF and LF words were extracted from Corpus Gesproken Nederlands (Spoken Dutch Corpus) available on www.tst.inl.nl and from the CELEX database on www.celex.mpi.nl. Each HF word with the context CəC- was matched with an LF word following a number of criteria: the consonants preceding the schwa and following the schwa were identical, the number of syllables was identical, the number of phonemes was matched as much as possible and stress placement was identical.

The corpus frequency counts may deviate from the frequency of the individual lexicon of the subjects. It is well-known in psycholinguistics that word frequency is related to naming latency, in the sense that HF words have shorter latencies than LF words (Oldfield & Wingfield 1965, Jescheniak &

Levelt 1994, among others). In order to ensure that the selection on the basis of frequency was justified, naming latencies were measured.

Methodology

Stimuli

Stimuli were selected on the basis of their frequency and the CWF. LF words had an occurrence of less than 10 per million and HF words occurred at least 100 times per million. Eight pairs of HF words and LF words were selected with possible target clusters /χl, bR, bl, vR/ and eight pairs of HF words and LF words were selected with impossible clusters {bχ, bd, bn, bz, dz, χm, χn, χv, χv}. In all, 32 critical stimuli were collected and 48 words were used as fillers, which were pseudo randomized.

Subjects

Twenty participants, mostly undergraduate and graduate students of Leiden University Linguistics Department, took part in the experiment. The group of subjects consisted of 13 females and 7 males. All of them were native speakers of Dutch. All subjects had normal or corrected-to-normal vision. The subjects took part voluntarily and were not paid for their participation.

Procedure

The subjects were individually recorded in an isolated booth. They received instructions from the experimenter and via a computer screen to quickly read aloud the words presented. The stimuli were made visible on a computer screen for 500 ms, followed by an interval of 1000 ms. The 80 stimuli were presented ten times in different orders. The order of the blocks was fully randomised. The naming latency was measured, making use of E-Prime standard 2.0. Each session was recorded with a *Sennheiser* MKH 416 directional condenser to a Marantz recorder. The computer was placed outside the booth. The experimenter was present to ensure that the experiment was carried out in the expected way. Schwa durations were measured in Praat (Boersma & Weenink 2001), measurements shorter than 10 ms were treated as deletions.

Results

Reaction times were measured to ensure that the selection of HF and LF words for individual subjects was justified. A t-test confirmed the justification of the selection except for one subject, so her responses were excluded from further investigation. Secondly, deletion of schwa was calculated. The number of deletions increases with repetition of the stimuli, but no higher tendency for deletion in HF words or well-formed clusters was detected.

Table 2. Percentages and number of schwa deletions based on the four stimuli groups

WordGroup	Deleted schwa		Undeleted schwa		Total	
	Perc	N	Perc	N	Perc	N
cluster ill-formed LF	11.1%	122	88.9%	982	100%	1104
cluster ill-formed HF	10.2%	113	89.8%	999	100%	1112
cluster well-formed LF	16.8%	187	83.2%	926	100%	1113
cluster well-formed HF	14.4%	140	85.6%	832	100%	972

So deletion does not seem to play a role: possibly deletion has to be treated as an extremely reduced form. To analyse the relative importance of the predictors, a regression test on WFC, frequency, repetition and the number of syllables showed a positive correlation ($R = 0.327$, $F=46.052$, $p=0.000$). The factors that significantly contribute to the variation effects were repetition ($t=-8.585$, $p<0.001$) and CWF ($t=-10.215$, $p<0.001$). Frequency effects were not found, probably due to overshadowing of repetition effects, since both are essentially a matter of activation. Therefore I conducted an ANOVA on the first block of stimuli only. The main effect of CWF was $F=32.437$, $p=0.000$. Although no main effect of frequency was attested, there was a significant interaction between CWF and frequency ($F=5.188$, $p=0.040$).

The diverging lines of the interaction diagram (Figure 1) show that the difference in duration between well-formed target onset clusters and ill-formed target onset clusters is relatively small in LF words, whereas it is larger for HF words. This means that HF words show more variation in schwa duration under the influence of phonotactic well-formedness of the cluster than LF words. Put differently, HF words are more sensitive to cluster well-formedness than LF words in reduction. Notice that it is not necessary for the schwa to be deleted for this interactional effect to occur.

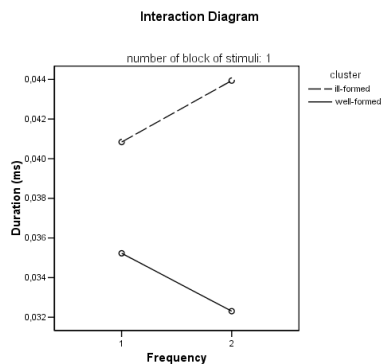


Figure 1. The interaction in pretonic schwa duration between frequency (1=LF, 2=HF) and well-formedness of the cluster.

Discussion and conclusion

The results of this experiment showed that word frequency interacts with phonological grammar such that high frequency words are more sensitive than low-frequency words to reduction under the influence of the well-formedness of the cluster that would remain after full deletion. This interaction must have consequences for phonological theory, since frequency effects and grammar are usually treated separately. Frequency effects have been modelled in Optimality Theory (see the work of Boersma 1998 seq. and Coetzee 2008), but not in direct connection to grammatical structure. Since frequency information is fundamentally lexical rather than grammatical, the interaction found in this experiment confirms that modelling on the basis of a grammar as well as a lexicon (as suggested by Ernestus 2009 and van de Weijer 2009, 2011) is on the right track.

References

- Boersma, P. 1998. *Functional Phonology*. Doctoral Disertation, University of Amsterdam.
- Boersma, P.J., Weenink, D. 2001. PRAAT [computer program].
- Booij, G. 1995. *The Phonology of Dutch*. Oxford, Clarendon Press.
- Coetzee, A.W. 2008. Phonological variation and lexical frequency. *Proceedings of NELS* 38.
- Ernestus, M. In press. Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua*.
- Jescheniak, J., Levelt, W. 1994. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition* 20, 824–843.
- Oldfield, R.C., Wingfield, A. 1965. Response latencies in naming objects. *The quarterly journal of experimental psychology* VII.1, 273-281.
- Oostendorp, M. van. 1998. Schwa in phonological Theory. *Glott International* 3(5), 3-8.
- Weijer, J. van de. 2009. Optimality Theory and Exemplar Theory. *Phonological Studies* 12, 117-124.
- Weijer, J. van de. submitted. Grammar as selection. *Lingua*.

Typology and spatial cognition in English, French and Greek: evidence from eye-tracking

Efstathia Soroli

Structures Formelles du Langage Lab, CNRS & University of Paris8, France

Abstract

Languages encode space in strikingly different ways (Talmy, 2000): *Satellite-framed* languages (e.g., English) lexicalize *Manner* in verb roots and express *Path* in satellites, whereas *Verb-framed* languages (e.g., French) lexicalize *Path* in verb roots, leaving *Manner* implicit or peripheral; other languages present *parallel* systems in which both *Verb-* and *Satellite-framed* structures are available (e.g., Greek). The present study investigates how speakers of three typologically different languages, English, French and Greek, performed a production task and allocated their visual attention while exploring and describing visual scenes involving motion events. The findings show that participants' verbalizations and eye-movement fixations differed substantially as a function of language-specific factors, arguing that typological constraints have a clear impact not only on linguistic but also on non-linguistic behaviours.

Key words: spatial cognition, language typology, production, eye-tracking method.

Introduction

Languages map semantic elements in very different ways when expressing motion. Talmy (2000) classifies languages into: *satellite-framed* (e.g., English) and *verb-framed* (e.g., French); the former are languages that lexicalize the *Manner* of motion in the verb and use satellites (e.g., particles) to express *Path* information within one compact structure (see example 1); the latter are languages that lexicalize *Path* in the verb stem, leaving *Manner* information implicit or placing it at the periphery of the sentence (e.g., gerunds/adverbials), (see example 2).

- (1) A man is running **into** a house.
MANNER PATH (S-framed pattern)
- (2) Il **entre** dans une maison en courant.
PATH MANNER
Lit. 'He is entering in[to] the house by running.' (V-framed pattern)

Another type of languages can equally manifest structures of both types. Greek displays such a mixed system, as illustrated in (3-4). However, for some authors, Greek is clearly a *verb-framed* language (Papafragou et al., 2006; Selimis, 2007), in which the V-pattern is dominant.

- (3) *Efyge trehontas (pros ta pano)*

- | | | | | |
|-----|---------------------------|---------------|------------------|--------------------|
| | PATH | MANNER | PATH | |
| | Left | running | (to-upwards) | |
| | 'He left running upwards' | | | (V-framed pattern) |
| (4) | <i>Etrekse</i> | <i>mesa</i> | <i>sto spiti</i> | |
| | MANNER | PATH | PATH | |
| | Ran | into | to-the house' | |
| | 'He ran into the house' | | | (S-framed pattern) |

An additional issue relevant to the typological classification of Greek is prefixation. Greek exhibits a set of preverbs of motion verbs (e.g., *ανεβάζω*, 'to up-put'; *εξέρχομαι*, 'to out-come'; *κατεβαίνω* 'to down-walk/go'; *ξεγλιστρώ* 'to out-slide') which function as satellites for *Motion*, *Path* and/or *Manner* verb roots. However, little is known about this phenomenon, which may raise important arguments regarding the real typological status of Greek. In the present study I examine preverbs of this type and the option of an additional *S-framed* sub-pattern that may exist as a result of verbal prefixes that share properties of satellites, as illustrated in (5).

- | | | | |
|-----|-----------------------|----------------------|-----------------------------|
| (5) | <i>Anevike</i> | <i>trehontas</i> | |
| | Ana [prefix]- | veno[Manner V-root] | treho[Manner V-gerund] |
| | PATH | MANNER/MOTION | MANNER |
| | up-walked/went | | running |
| | 'He ascended running' | | (S-framed prefixed pattern) |

Such striking cross-linguistic differences, apart from their typological interest, are significant for the study of the relationship between language and cognition, and contribute to the debate opposing *universalist* approaches—according to which spatial cognition is based on universal perceptual and cognitive processes that are independent from language-specific properties—and *relativity* approaches—according to which language-specific factors affect how speakers construct spatial representations. In the context of this debate and while most studies showing language effects have been based on language use (Hickmann, et al. 2009; Slobin, 2006), the present study analyses verbal and non-verbal responses of speakers of typologically different languages (cf. Soroli & Hickmann, 2010a) in order to determine the extent to which language properties influence different measures of cognition.

Method

The present study investigated how 42 native speakers of English, French and Greek (14 per language) performed a verbal task and allocated their visual attention while exploring and describing visual scenes (clips)

involving events of voluntary motion. The analysis of the verbal measures examined two aspects of the responses: the types of information selected for expression (*focus*) and the means whereby this information was expressed (*locus*). In Greek (see Soroli & Hickmann 2010b; Hickmann et al., submitted), the data were coded twice (*Verb-* vs. *Satellite-*coding). *V-coding* did not differentiate prefixed and plain verb forms, while *S-coding* decomposed prefixed verb forms into a verbal root and a satellite-like verbal prefix. As a result, verbal prefixes were treated as part of the verb in *V-coding*, but not in *S-coding*. The analysis of eye-movements examined the number, length and timeline of the fixations in specific areas of interest (*Path* vs. *Path±Manner*) in the clips.

Results and discussion

Subjects' verbalizations differed substantially as a function of language-specific factors. English speakers used compact structures that expressed both *Manner* and *Path* information and in which they systematically encoded *Manner* in verbs and *Path* in other devices. In contrast, French speakers mostly focused on *Path* information that was expressed in the verb, and either provided less information about *Manner* in the periphery of the sentence or did not use any other devices in the verbal network to express this information. As for Greek, it seems to present a third pattern, since some utterances seem to be *V-framed*, others are *S-framed*, and a third group seem to belong to an *S-framed* sub-system comprising *Path* prefixes. These patterns directly follow from language-specific factors that typologically differentiate English, French and Greek as *Satellite-*, *Verb-* and *Parallel-framed* languages respectively.

Turning now to eye-movements, although all speakers allocated more attention to *Path* overall, their focus also varied across language groups in all measures. More specifically, French speakers focused their gaze mostly on *Path* areas and this from stimulus onset until the end of the processing timeline. In contrast, the gazes of English and Greek speakers showed no preference for *Path* or *Path±Manner* areas of interest, with the exception of brief *Path* preferences at the onset of the stimuli for English speakers.

Conclusion

Overall the verbal and non-verbal findings point to differences in behavioural patterns as a function of typological language properties. In conclusion, we argue that typological constraints have a clear impact on linguistic behaviour, but also on non-linguistic behaviour. Such results contribute to current debates concerning the language-thought interface. In particular, linguistic analyses based on controlled methodologies, such as the one that was used in the present study, can yield comparable databases

across a number of languages, thereby making it possible to put forth fine-grained and realistic hypotheses concerning the cognitive implications of typological properties.

Acknowledgements

I wish to thank Maya Hickmann for kindly providing her stimuli for this study and precious advice. Eleni Valma, Stathis Selimis, Zaphira Rohde, Christina Papadimitraki and Annemarie Verkerk, for much discussion and feedback on Greek. I am especially grateful to Lina Soroli, Sofien Guedanna, Maria Papageorgiou, Yiannis Vlassopoulos and Jean-René Borelly for help in experiment preparation and much inspiring discussion.

References

- Hickmann, M., Tarrane, P., Bonnet, Ph. 2009. Motion in first language acquisition: Manner and Path in French and in English. *Journal of Child Language*, Vol 36, Issue 04, 705-741.
- Hickmann, M., Hendriks, H., Soroli, E., Iakovleva, T., Ji, Y. submitted. Space and language typology : encoding motion across languages.
- Papafragou, A., Massey, C., Gleitman, L. 2006. When English proposes what Greek presupposes: The linguistic encoding of motion events. *Cognition*, 98, B75–B87.
- Selimis, E. 2007. Γλωσσική κωδικοποίηση του εννοιολογικού πεδίου της κίνησης: κυριολεξία και μεταφορά στα ελληνικά παιδιών και ενηλίκων ('Linguistic coding of the concept of motion: Literal and metaphorical expressions in adult and child Greek'). Unpublished Doctoral dissertation, University of Athens.
- Slobin, D.I. 2006. What makes manner of motion salient? Explorations in linguistic typology, discourse, and cognition. In M. Hickmann, S. Robert (Eds.), *Space across languages: linguistic systems and cognitive categories*, 59–81. Amsterdam: John Benjamins.
- Soroli, E., Hickmann, M. 2010a. Language and spatial representations in French and in English : evidence from eye-movements. In G. Marotta, A. Lenci, L. Meini & F. Rovai (Eds.), *Space in Language*, 581-597. Pisa : Editrice Testi Scientifici.
- Soroli, E., Hickmann, M. 2010b. Crosslinguistic spatial cognition: exploring visuospatial thinking and speaking. Presented at the Summer Institute in Cognitive Sciences 2010: The Origins of Language. Montreal, Canada, June 21-30, 2010
- Talmy, L. 2000. *Toward a cognitive semantics. Volume 1: Concept structuring systems. Volume 2: Typology and process in concept structuring*. Cambridge, MA: MIT Press.

The processing of asymmetric and symmetric sentential conjunction

Ellen Thompson, Javier Collado, Maria Omana, Amanda Yousuf-Little
Linguistics Program, Department of English, Florida International University,
U.S.A.

Abstract

In this study, we examine the predictions for processing of a syntactically articulated theory of the distinction among different interpretations of clausal 'and'. Bjorkman (2010) claims that symmetric 'and' interpretations involve coordination of CPs; these are logical interpretations. Asymmetric interpretations of 'and' involve conjunction of TPs; these are temporal and causal. If the processor is guided by structural considerations, we predict a possible two-way split in the processing costs of these structures. Therefore, this research examines the processing time involved in sentences interpreted as: (i) temporal, (ii) causal, and (iii) logical, versus the distinctions of (i) asymmetric (TP structure), and (ii) symmetric (CP structure). We find that structures involving symmetric 'and' involve longer processing times than those of asymmetric, causal 'and', and although the processing times of structures with logical 'and' are longer than those with temporal 'and', this distinction does not approach statistical significance.

Key words: coordination, conjunction, processing time

Introduction

Within Generative Grammar, multiple analyses of the distinct interpretations associated with the clausal conjunct 'and' have been developed (Culicover 1970; Posner 1980). This paper focuses on the logical interpretation of 'and' illustrated in (1a), the temporal interpretation shown in (1b), and the causal meaning, as in (1c):

- (1) a. Water freezes at 0°C and ethanol freezes at -114°C.
- b. The lights came on and the singer stepped onto the stage.
- c. The lights were off and I couldn't see.

It has been noted that logical 'and' is symmetric; it allows a reversal of the two conjuncts with a maintenance of meaning; (2a) is equivalent to (1a):

- (2) a. Ethanol freezes at -114°C and water freezes at 0°C.

In contrast, temporal and causal 'and' do not permit reversal of the two conjuncts with the same meaning, ((3a) versus (1b) and (3b) versus (1c)) and these uses are therefore characterized as asymmetric.

- (3) a. The singer stepped onto the stage and the lights came on.
b. I couldn't see and the lights were off.

In contrast to earlier claims that it is pragmatic and discourse factors which determine the interpretation of clausal 'and', Bjorkman (2010) argues that the difference between symmetric and asymmetric 'and' is semantic. She claims that this semantic distinction is reflected in the syntax of the conjunction structures: symmetric coordination involves conjunction of CP structures, whereas asymmetric coordination involves conjunction of TP structures.

We investigate the prediction for processing of such a syntactically articulated theory. It is predicted by this approach that the processing cost associated with the comprehension and production of these two interpretations of clausal 'and' is different. Processing of asymmetric conjunction structures involve conjunction of TPs, and thus less structure than the processing of symmetric 'and' structures, which require the conjunction of necessarily larger structures, CPs. Assuming that processing cost is associated with syntactic structure that is phonologically covert, as well as with phonologically overt material, the difference in processing time between these two distinct types of coordination is predicted to be measurable.

Methods

We tested this hypothesis with eight adult monolingual English speakers using the Rapid Serial Visual Presentation (RSVP) method via a PowerPoint presentation. Subjects are displayed sixty sentences in total, composed of thirty filler sentences as well as ten causal, ten temporal, and ten logical sentences.

Sentences are displayed a single word at a time for 800 ms. Participants silently repeat each word to themselves, without labial movement. At the end of each sentence, participants are prompted by a marker "X" and a tone. This indicates the completion of the sentence and prompts the subject to reproduce the target sentence. The processing time of each coordinate structure is processed using WavePad Sound Editor by measuring the distance between the beginning of the sounded prompt and the completion of the participants' utterance.

Results

A one way repeated measure ANOVA was carried out on the processing times for the three coordinate structures. Post hoc tests were carried out using Fisher's LSD pair wise comparison at the 5% significance level. There was a significant difference in coordinate structures, $F(2,14) = 6.86$, $p < .017$. Fisher's tests indicated that the mean processing time for causal ($M = 3.48$

sec) was significantly less than temporal ($M = 3.86$) and logical ($M = 4.23$). A significant difference was not found between temporal and logical *and* structures. Refer to Table 1, Comparison of Production Times of Coordinate Structures.

Table 1. Comparison of Production Times of Coordinate Structures

	Z	Asymp. Sig. (2-tailed)
temporal - causal	-1.960 ^a	.050
logical - causal	-2.100 ^a	.036
logical - temporal	-1.540 ^a	.123

Due to the presence of several high processing times, the nonparametric Friedman's test was performed to confirm the ANOVA results; similar results were found. There was a significant difference among the three coordinate structures, $X^2(2, N=8) = 7.00$, $p < .03$. The causal structure had significantly less processing time than the temporal and logical structures, while temporal and logical did not differ between them.

Figure 1 illustrates the mean production time of each coordinate structure, demonstrating the higher processing time associated with logical and temporal conjunctions in contrast to causal *and* conjunctions.

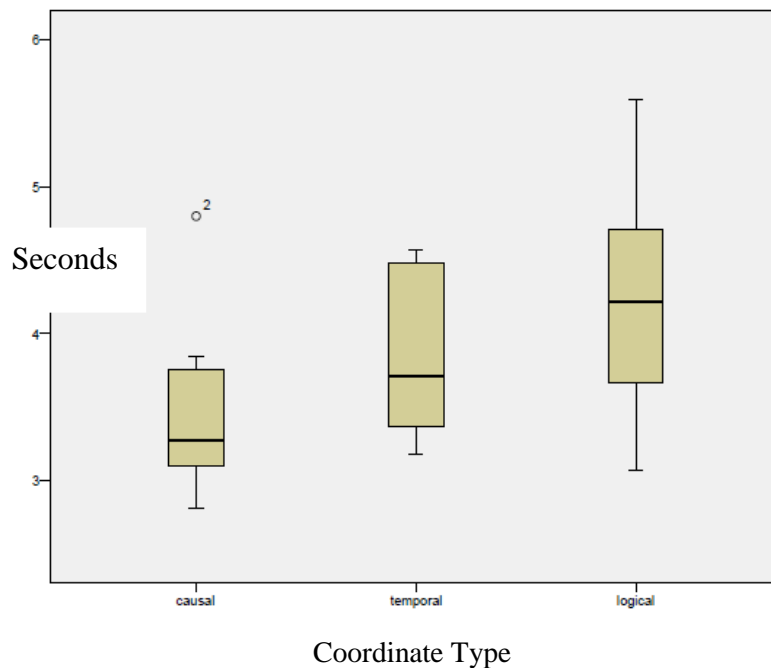


Figure 1. Mean Production Times of Coordinate Structures

Discussion

Recall that the prediction of a two-way distinction between the syntactic and semantic structures of symmetric and asymmetric 'and' is that logical 'and' structures should have a higher processing time than causal and temporal 'and' structures. The data from this study provide partial confirmation of these predictions: structures involving symmetric 'and' involve longer processing times than those of asymmetric, causal 'and'. However, although it is the case that the processing times of structures with logical 'and' are longer than those with temporal 'and', this distinction does not approach statistical significance. This effect is not predicted, and the implications of this effect for the theory of coordination are to be discussed.

Acknowledgements

We would like to extend our gratitude for financial support for this research to the Graduate Student Funding Committee, the College of Arts and Sciences, and the English Department at Florida International University. We also want to thank for their support and comments: Antonis Botinis, Gisela Casines, Marcela Depiante, Maureen Donnelly, Ana Luszczynska, James Sutton, and Kezia Walker.

References

- Blakemore, D., Carston, R.. 2005. The pragmatics of sentential coordination with 'and'. *Lingua* 115.
- Bjorkman, B. 2010. A Syntactic Correlate of a Semantic Asymmetry, paper presented at NELS 41, October.
- Culicover, P.W. 1970. One more can of beer. *Linguistic Inquiry* 1, 366–369.
- Culicover, P.W., R.S. Jackendoff. 1997. Semantic subordination despite syntactic coordination. *Linguistic Inquiry* 28, 195–217.
- Posner, R. 1980. Semantics and pragmatics of sentence connectives in natural language. *Speech act theory and pragmatics* 168–203.
- Progovac, L. 1998. Structure for coordination. Part II. *Glott International* 3, 3–9.
- Schmerling, S.F. 1975. Asymmetric conjunction and rules of conversation. In Cole, P., Morgan, J.L. (eds.), *Syntax and semantics volume 3, Speech Acts*, 211–232. Academic Press.
- Txurruka, I.G. 2003. The natural language conjunction 'and'. *Linguistics and Philosophy* 26:255–285.

Identifications of speaker-ethnicity: Attribution accuracy in changeable settings

Richard Todd

Speech and Hearing Research Group, Department of Computer Science, University of Sheffield, UK

Abstract

Several studies have considered the auditory identification of foreign-accented/non-native speech. Here, a finer-grained alternative to the traditional definition of nativeness is used. The approach eliminates the shortcomings of selecting ‘native’ speech based on the usual criterion of birthplace alone, yet accommodates complexities which arise when dealing with speakers from countries which are historically diverse, ethnically. Using the foregoing construct, this study examines the potential for listeners to accurately group different speaker-types, with respect to their ethnicity, in low/high quality transmission conditions. The findings of the Ethnic Group Attribution (EGA) task confirms overall human competence. The work’s large number of participants ($n = 120$) brings added generalisability to smaller-scale studies. It furthermore, allows a better understanding of contextualised performance, gender-wise.

Key words: Ethnic Group Attribution (EGA); Ethnicity; Nativeness; Identification; Perception.

Introduction

Listeners’ ability to identify or characterise speakers in terms of their foreign-accentedness or ethnicity has attracted interest in a number of ways. This is evidenced by the investigations of Arslan and Hansen 1996; Flege 1988; Sebba 1993; Todd 1998; Walton & Orlikoff 1986, for example. It becomes clear from such works that delineating nativeness or ascribing ethnicity to voices heard using the same language is a non-trivial task. Furthermore, performing the latter challenge — Ethnic Group Attribution (EGA) — requires listeners to additionally attend to features associated with *speaker-types*, or *-groups*, rather than highly idiosyncratic qualities which, in forensic-phonetic situations, usually promote fixation on one speaker.

It is well-understood that factors such as locality, social networks/peer affiliations, and language familiarity may variously influence speech perception. Clopper et al. 2006; Kerswill et al. 2008 encapsulate the general affects of speaker|listener locale and social networks on identification. Clopper 2004; Sullivan & Schlichting 2000 further illustrate any benefits of on prior linguistic awareness. Notwithstanding the foregoing however, implicit ambiguities that may exist when speakers from countries/regions having a long-standing history of ethnic diversity (like, England or the USA) are considered for study. The traditional native|non-native criterion,

operating purely on lines of birthplace, appears to serve us well, initially. Inadequacies of the binary approach arise however, on realising speakers of variable descent (say, Italian, Ukrainian, etc.) may also be born and raised in the same locality as others who are ‘native’ via traditional definition.

The predicament means differing speaker-types may unwittingly populate a stimulus group. Devising a more refined framework allows speakers having shared overarching features to be more clearly disambiguated, especially for the sake of others wishing to further develop acoustic and auditory studies alike (cf. Kerswill et al. 2008). Figure 1, below, illustrates the how speaker-nativeness is nuanced. From both socio- and forensic-phonetic standpoints, it is believed this finer-grained approach may, respectively, permit greater openings for study and debate.

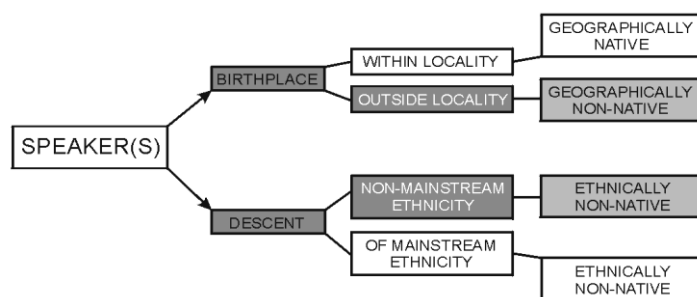


Figure 1. A new framework of Nativeness considers speakers’ ethnicity in addition to geographical origins, to allow finer-grained distinctions.

In this work, nativeness is vitally considered in the above terms, given that speaker-ethnicity and its identification is are central themes.

The Study

Without doubt, work on the perceptions of non-native speech from an L2 perspective continues to steadily grow. Attention however, still seems to bypass the issue of whether listeners can reliably group speakers in terms of their respective ethnicities, especially if all (1) inhabit a common locality; (2) have routine/innate familiarity with those speech norms; and, (3) were raised with some other language/system(s) also in use (L_n).

Here, EGA potential in high- and low-quality listening conditions is examined while encompassing the three aforementioned points.

Participants and Method

Various negative constraints mean studies related to either the identification or perception speaker-accentedness or -ethnicity may feature a relatively modest number of participants. To illustrate, Yuang et al. (2010) considered the attributions of just 3 participants. However, Clopper (2004) and Kerswill et al. (2008) improve generalisations further (where, respectively, speaker types = 6 and 4; $n = 49$ and 68 subjects, respectively, from the same locale). Despite this, both of the latter works are problematic, if seeking to precisely determine the ethnicity of any speech considered.

An even greater number of participants were considered in this investigation ($n = 120$). All were adults (mean age = 34.4; male = 52; female = 68). While, overall, each speaker-group would be of the participants' locale, non-native ethnicities were South Asian, Caribbean, and East Asian (total $n = 45$). Research volunteers had to consider such speech (being lexically identical) while presented in a High-Quality setting (with face unseen; HQ), and Low-Quality setting (i.e., via telephone; LQ).

Results

Table 1. Summarised attribution potential (overall, and gender-wise).

Settings considered	Overall Ratings [†]			Female Ratings [†]			Male Ratings [†]		
	Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
Face Unseen	2.08 (sd 0.8)	1	4	2.22 (sd 0.83)	1	4	1.88 (sd 0.73)	1	3
Telephone	2.28 (sd 0.86)	1	5	2.43 (sd 0.85)	1	5	2.08 (sd 0.84)	1	4

[†] rating 1 = all the time; 2 = frequently; 3 = occasionally; 4 = hardly ever; 5 = virtually never

Besides gender-wise participant means, Table 1, above, shows highest/lowest ratings which, practically, were analogous to the accuracy and confidence of EG as being made. In terms of descending competence/accuracy, rating '1' = 75-100%; '2' = 50-75%; '3' = 25-50%; '4' = 10-25%; and '5' = 0-10%. Males yielded higher, yet less varied, performance ratings, suggesting the more assured identifications of speaker-ethnicity. Significant gender differences in performance outcomes existed. T-test p -values moved from $p = 0.022$ in the HQ (face unseen) setting, to $p = 0.027$ when making comparisons for the least reliable LQ (telephone) setting. Latterly, male and female EGA potentials dropped, to equal 73% and 64.25%, respectively (68%, overall). Previously this was 73% overall, where males showed 78% accuracy; females claimed 69.5%.

Discussion and Conclusions

The aim of the study was to bring more clarity and attention to an increasingly inadequate notion of speaker-nativeness, plus the under-researched area regarding the identifiability of speaker-ethnicity.

Despite limited space, this paper initially presents a framework which challenges, and may influence, current assumptions of nativeness. By featuring two stimulus qualities, the study allows better understanding of how extrinsic matters may influence listener perception, and thus, the potential outcome of attributions. Unlike other studies, participants contemplated speakers of a common locality, all producing like utterances, yet clearly differing in terms of direct ethnic descent. Results of this large-scale work in brief show listeners may, overall, claim '*frequently*' accurate performance in favourable conditions, despite gender differences.

References

- Arslan, L.M., Hansen, J. 1996. Foreign Accent Classification in American English. *Speech Communication*, 18, 353-367.
- Flege, J.E. 1988. Factors Affecting Degree of Perceived Foreign Accent in English Sentences. *Journal of the Acoustical Society of America*, 84, 70-79.
- Sebba, M. (1993) *London Jamaican: Language Systems in Interaction*. London: Longman.
- Sullivan, K., Schlichting, F. (2000) 'Speaker Discrimination in a Foreign Language: First Language Environment, Second Language Learners. *International Journal of Speech Language and the Law*, 7, 1, 95-112.
- Todd, R. 1998. Auditory Perception and Ethnic Group Attribution of Unknown Voices: Assessing the Robustness of Experienced Listeners' Ratings when Confronted with Non-Native but Proficient English Speech. *Proceedings of The Institute of Acoustics*, 20, 6, 343-350.
- Todd, R. 2002. Speaker-Ethnicity: Attributions Based on the Use of Prosodic Cues. *Proceedings of the 1st International Conference on Speech Prosody*, Speech Prosody 2002, 663-666.
- Walton, Orlikoff. 1986. Speaker and Race Identification from Acoustic Cues in the Vocal Signal. *Journal of Speech and Hearing Research*, 37, 738-745.
- Clopper, C. G. (2004). *Linguistic Experience and the Perceptual Classification of Dialect Variation*. Doctoral dissertation, Indiana University.
- Clopper, C. G., Levi, S. V., & Pisoni, D. B. (2006). Perceptual similarity of regional varieties of American English. *Journal of the Acoustical Society of America*, 119, 566-574.
- Yuan, J., Jiang, Y., Song, Z., 2010. Perception of foreign accent in spontaneous L2 English speech, *Proceedings of Speech Prosody 2010*, 100884:1-4.
- Kerswill, P., Torgersen, E., Khan, A. and Fox, S. Perceiving ethnicity and place in Multicultural London English. Paper presented at NWAV37, Rice University, Houston, November 2008.

Participle Agreement and Clitic Omission

Vicenç Torrens

Facultad de Psicología, National University of Distance Learning, Spain

Abstract

Object clitic omission has been considered to be a general feature of child grammar. In this paper we show that while omission is high in young Catalan-speaking children, it is very low in Spanish-speaking children. This difference can be attributed to the availability of participle agreement when objects cliticise. I present a research where we test, through a grammaticality judgement task, the sensitivity of Catalan and Spanish children to participle agreement and find a statistically significant difference between the two languages.

Key words: clitics, omission, participle agreement.

Theoretical Background

In the literature of language development, some analyses have been put forward to account for the phenomenon known as object clitic omission (see e.g. Guasti, 1993/1994, Hamann, Rizzi and Frauenfelder, 1996; Avram 2000, Schaeffer 2000, Jakubowicz et al. 1997). In this paper we challenge the idea that clitic omission is a universal feature of all child grammars, and show that it holds of only a subset. While object clitic omission is found in languages such as French and Italian, we demonstrate that it is found to a significantly lesser extent in other languages, such as Spanish, and we argue that there is a correlation between object clitic omission and participle agreement. This correlation can be accounted for under Wexler's (1998) Unique Checking Constraint.

The hypothesis we entertain is that in child grammar object clitic omission results from the interaction of the properties of some grammars with a maturational principle, the Unique Checking Constraint (UCC) (Wexler 1998). There is nothing that singles out clitics and their acquisition; rather, there is a more general constraint in the child's grammar, the UCC, which grants, for some languages, optional clitic omission, together with other features of child grammar, the most well-known of which is the optional infinitive stage. The Unique Checking Constraint, which acts in conjunction with Minimise Violations, is formulated as follows:

Unique Checking Constraint: the D-feature of DP can only check against one functional feature.

Minimise Violations: Given an LF, choose a numeration the derivation of which violates as few grammatical properties as possible. If two numerations are both minimal violators, either one may be chosen.

The UCC limits the derivations that are possible for the child to compute; if the target language of the child requires a derivation with double checking of uninterpretable features by a certain grammatical constituent, the child grammar will render that derivation a violation of the UCC. However, should that derivation compete with another derivation also involving a grammaticality violation, both derivations would be, by Minimise Violations, equally bad for the child's grammar, and would thus be expected to occur optionally.

Since the different omission rate between Catalan and Spanish is attributed to the existence/lack of participle agreement in these languages, the assumption is made that participle agreement is a feature of grammar from early on. Agreement has been found to be an early acquisition in other domains such as subject-verb agreement. However, the only direct evidence we have so far for the presence of participle agreement in child grammar comes from the production of participle agreement, which is not found for all children (nor for all adults) in standard Catalan. Since we claim that the mechanisms underlying participle agreement are the same for all children, whether participle agreement is overt or not, it is relevant to give independent evidence of the existence of participle agreement in child Catalan.

Experiment

Procedure

In order to test the availability of participle agreement in Catalan and Spanish, we ran a grammaticality judgement task. We tested 40 Catalan-speaking children (2;2,12-6;0,25) and 40 Spanish-speaking children (2;5,13-4;10,8) and 10 control adults from the same areas where the first experiment was carried out.

The task was designed to see whether participle agreement could be a constitutive feature of child Catalan grammar despite its optionality, and how it differed from Spanish. The experimental items included sentences with participle agreement, sentences without participle agreement and, only for Catalan, sentences with unmatched participle agreement. The grammaticality judgement task consisted of nine sentences for Catalan and eight sentences for Spanish. A grammaticality judgement task with subject-verb agreement mismatches was introduced as control, and children who could not detect them at all were excluded as unable to perform a grammaticality judgement task.

Children were first introduced to the task by means of ill formed word order sequences. Then, after they had been familiarized with these examples, control and experimental items were presented to them in pseudorandom order. The experimenter showed the child short stories consisting of two

pictures portraying an event; the story was described by two sentences, the first of which provided the antecedent for the clitic in the second. The second sentence was pronounced clearly and the child was asked to judge whether it was a good sentence for him/her or not.

Results

For children, the rates of acceptance appear in Tables 1 and 2 (where A agreement, N no agreement, SA subject-verb agreement). The adult controls provided the expected responses; in Spanish A and ungrammatical SA were rejected and only N was accepted in all cases.

Table 1 Rates of acceptance in Catalan.

	A	N	*SA
2-year-olds i	9/9	9/9	0/3
3-year-olds	27/27	27/27	3/18
4-year-olds	24/30	28/30	0/20

Table 2. Rates of acceptance in Spanish.

	*A	N	*SA
2-year-olds i	8/15	14/15	0/10
3-year-olds	2/30	30/30	0/20
4-year-olds	0/30	30/30	0/20

The results indicate that adult-like performance is achieved at three for Spanish and at five for Catalan. If we look at the acceptance rate of matched participle agreement, Catalan-speaking children are target-like at age three also. The two languages display very different patterns: while in Spanish there is early rejection of participle agreement and this develops steadily, in Catalan there is overacceptance of mismatched agreement (differences statistically significant between Catalan and Spanish for the agreement condition, for all ages ($P < 0.001$)). This result witnesses the child's sensitivity to participle agreement, and also gives empirical ground to a formerly unverified assumption: the availability of participle agreement in child grammar from early records – critical for the Wexler et al. analysis of clitic omission to hold.

Conclusion

The results of our elicitation experiment allows us to conclude that object clitic omission is a feature of the early stage of grammatical development, which is nevertheless not found universally. This crosslinguistic variation can be accounted for under a maturational principle, the UCC, which sets limits on the checking operations in a derivation. Some languages' object

clitics – e.g. those of Spanish – do not require the application of double-checking operations and are therefore adult-like from the earliest productions. Some other object clitics – e.g. those of Catalan – require double checking operations due to an additional participle agreement operation. There is one further consequence of our study regarding the optional character of participle agreement in the variety of Catalan tested, when compared to Italian: the optionality of participle agreement appears to be inconsequential. In the grammaticality judgement task, we provide independent evidence that child grammar is sensitive to the participle agreement feature in the target language, given that the results for Catalan sharply contrast with those for Spanish.

References

- Avram, L. 2000. Clitic Omission in Child Romanian. Psycholinguistic Colloquium. University of Potsdam.
- Guasti, M.T. 1993/1994. Verb syntax in Italian child grammar: finite and non-finite verbs. *Language Acquisition* 3, 1-40.. Place, Publisher.
- Hamann, C., Rizzi, L., Frauenfelder, U. 1996. On the Acquisition of the Pronoun System in French. In Clahsen, H. (ed.), *Generative Perspectives on Language Acquisition*. Amsterdam, John Benjamins.
- Jackubowicz, C. 1997, Müller, N., Biemer, B., Rigaut, C 1997. The Case of Subject and Object Omission in French and German. In Hughes, E., Hughes, M, Greenhill, A. (eds.), *Boston University Conference on Language Development Proceedings*, 21, Somerville, MA, Cascadilla Press.
- Schaeffer, J. 2000. The Acquisition of Direct Object Scrambling and Clitic Placement. Amsterdam, John Benjamins.
- Wexler, K. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106, 23-79.

High functioning autism and prosody of sentence types in Greek: A case study

Maria Tripolitou¹, Anthi Chaida^{2,3}

¹Queen Margaret University of Edinburgh, UK

²Department of Speech & Language Therapy, AKMI Metropolitan College, Greece

³Laboratory of Phonetics & Computational Linguistics, University of Athens, Greece

Abstract

The study focuses on the prosodic characteristics of sentence types in Greek, based on data by an adult with high functioning autism (HFA). It investigates the effects of the disorder on tonal and temporal features of 2 basic sentence types. A production experiment was carried out, based on SVO sentences produced as statements and polar questions, with no other difference than their prosodic realisation. The HFA findings were compared to relevant findings for normal speech and showed: (a) very similar tonal patterns, (b) more expanded tonal, and (c) longer utterances.

Key words: prosody, intonation, sentence types, high functioning autism, Greek.

Introduction

The present study reports on the prosodic characteristics of sentence types in Greek, based on a case study of an adult with high functioning autism (HFA). It aims to investigate the effects of HFA on tonal and temporal features of 2 basic sentence types: statement and polar question. The findings were compared to relevant findings for normal speech.

High functioning autism and prosody

Communicative dysfunction is a core symptoms of autistic syndromes in individuals who speak, although there is no significant delay in language development of people with HFA. Expressive language is often stereotyped. Vocal characteristics can be very different in these individuals. Differences reported include monotonic or machine-like intonation, deficits in the use of pitch and control of volume, deficiencies in vocal quality, and use of aberrant stress patterns. Productions may sound rather flat, due to monotone, robotic phrasing, artificial sounding and inflection patterns (Happé, 1996). When these differences are present, they tend to be persistent and show little change over time (Shriberg et al., 2001). It has been found that HFA speakers may produce utterances without a terminal pitch contour, and may have either very narrow or very wide tonal ranges, also lacking correlation between frequency and intensity (Baltaxe et al., 1984). However, HFA it has been reported that tonal boundaries may be adequately used to locate ends of utterances (Fine et al., 1991). Wide pitch variations have been also noted. With regards to temporal features, HFA speakers have been reported to produce longer utterances with high variability (Baltaxe et al., 1984).

Prosodic deficits in HFA, however, have not been universally reported, in general, while -to our knowledge- there is no relevant research concerning the Greek.

Prosodic patterns of sentence types in Greek

In Greek, in the absence of all lexical information, the distinction between sentence types is assumed to depend on the prosodic structure. Sentence types are associated with local and global tonal structures. Questions are mostly associated with higher tonal register, less tonal declination and a final tonal rise (e.g. Botinis, 1998). Studies in Greek prosody have revealed several tonal features that distinguish each sentence type (e.g. Baltazani, 2007; Botinis, 1998; Chaida, 2007, 2010; see Fig. 1).



Sentence Type	Tonal structure	Boundary
STATEMENT		Low
POLAR QUESTION		Rise-Fall

Figure 1. Stylized tonal structures for statement and polar question in Greek (adapted from Chaida, 2007, 2010).

Methodology

Hereby, we present a production experiment based on a case study of a 43-year-old female adult with HFA, native speaker of Greek, with the ability to read. The speaker produced three SVO sentences, as statements and polar questions, based only on prosodic variations, without any other lexical, morphological or syntactical information given. The phonetic material consisted of a corpus of 60 sentences in total (10 repetitions X 3 randomised repetitions X 2 sentence types). Three F0 measurements were taken per syllable, while the length and the pitch range of each utterance were also noted down. The recorded speech samples were analyzed and assessed in contrast to the results of a control group of 5 female normal speakers (Chaida, 2010).

Results

HFA findings are presented in comparison to normal speakers' findings in Table 1 and Figures 2-4.

As shown in Table 1, the **pitch range** is higher in statement than in question, both in HFA and normal speech. It should be noted that the pitch range in HFA is significantly higher than the pitch range in normal speech ($t(df\ 2)=7.28, p=0.01$).

Table 1. Mean pitch range (in Hz) of statements and questions in HFA and normal speech (female speakers).

Speakers	Sentence Type	Min.	Max.	Pitch Range
HFA	STATEMENT	106	338	232
	QUESTION	107	316	209
Normal	STATEMENT	173	254	81
	QUESTION	162	269	107

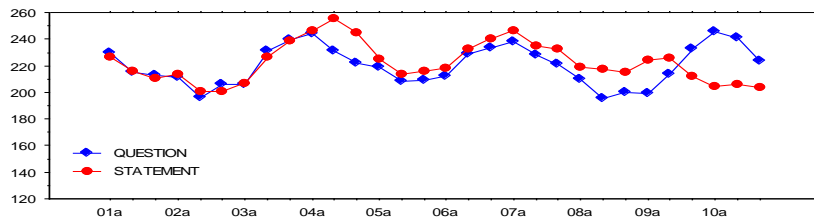


Figure 2. Mean tonal patterns (Hz) of statement and question by an HFA female speaker.

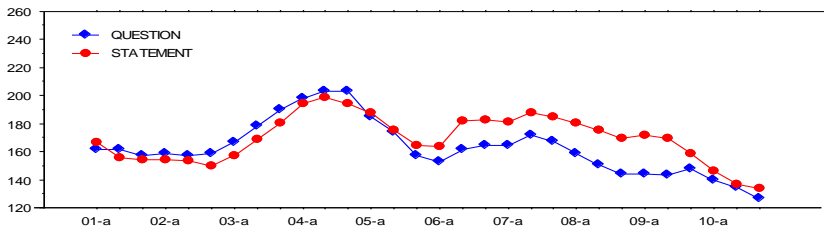


Figure 3. Mean tonal patterns (Hz) of statement and question by 5 normal female speakers.

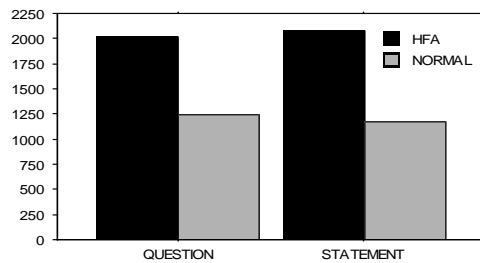


Figure 4. Mean duration (in ms) of statements and questions in HFA normal speech (female speakers).

As presented in Fig. 2-3, regarding **intonation**, the tonal pattern for *statements* shows very few differences: the utterance tonal offset in HFA is higher than the one in normal productions. Regarding *questions* in normal speech, the peak of the rise-fall of the tonal boundary is aligned with the

final stressed syllable of the utterance, i.e. the pre-final syllable of the utterance. In HFA, the peak of the tonal boundary rise-fall in questions is aligned with the final syllable of the utterance, while it is a lot higher than in normal speech, leading to a higher utterance tonal offset. The difference between HFA and normal tonal patterns is statistically significant for questions ($t(df\ 928)=2.03$, $p=0.01$), but not for statements ($t(df\ 928)=1.55$, $p=0.11$).

Sentence type differences with regards to **duration**, are of no significance both for HFA and normal productions, as shown in Fig. 4. However, the duration difference between HFA and normal speech is highly significant, both for questions (771,80 ms; $p<0.0001$, at 0.05 level) and for statements (900,95 ms; $p<0.0001$, at 0.05 level).

Conclusions

The results give a clear picture that HFA sentence type production. HFA findings showed very similar tonal characteristics to normal speech, concerning the intonation of statements and polar questions in Greek. However, the tonal range of the HFA speaker is significantly more limited than that reported for normal female speakers. With regards to temporal characteristics, HFA productions were found to be significantly longer than the normal ones. The prosodic deficits frequently attributed to people with autistic syndromes reside primarily in the pragmatic and affective aspects of prosody, with grammatical aspects relatively spared (Shriberg et al., 2001). In general, HFA intonation findings of the present study revealed that there is a clear distinction between sentence types, indicating that prosody may not be considered as a major communicational constraint for people with HFA.

References

- Baltaxe C.A.M., Simmons J.Q., Zee E. 1984. Intonation patterns in normal, autistic and aphasic children. In Cohen A. & Van de Broecke M. (eds), *Proc. of the 10th International Congress of Phonetic Sciences*, Dordrecht: Foris.
- Botinis, A. 1998. Intonation in Greek. In Hirst, D. and Di Cristo, A. (eds), *Intonation Systems: A Survey of Twenty Languages*. CUP, 288-310.
- Chaida A. 2007. Tonal Structures of Complex Sentences in Greek. In *Proc. of the 8th International Conference on Greek Linguistics*, Ioannina, Greece.
- Chaida A. 2010. [in Greek] Production and perception of sentence type intonation in Greek. PhD Thesis, University of Athens, Greece.
- Fine J., Bartolucci G., Ginsberg G., Szatmari P. 1991. The use of intonation to communicate in pervasive developmental disorders. *Journal of Child Psychology and Psychiatry*, 32, 777-782.
- Happé F. 1996. *Autism*. London: UCL Press.
- Shriberg L.D., Paul R., McSweeney J., Klin A., Cohen D.J., Volkmar F.R. 2001. Speech and Prosody characteristics of adolescents and adults with high functioning autism and Asperger syndrome. *and Hearing Research*, 44, 1097-1115.

Metacommunicative devices in spoken discourse as part of processing distributed cognitive tasks

Ilya Utekhin, Tatiana Chernigovskaya

Department of General Linguistics, St.Petersburg State University, Russia

Abstract

Types of utterances that belong to meta-levels of activity have been singled out in spoken task-focused dialogues between subjects accomplishing a matching task with no visual contact. The types include activity management and planning, global and local (e.g. signals of activity phase); explicit evaluation of the state of joint project, of one's own state or partner's state; retrospective references and accounts; meta-communicative utterances for conversation management; backchannel response: "continuers"; collaborative completion, echoic quotation; repair (self-repair, other-repair); meta-linguistic utterances; meta-communicative utterances having to do with rapport, etiquette, joking, etc. Differences between normal and schizophrenic subjects have been shown in meta-communicative activity supposed to be linked to the management of distributed cognition and creation of shared representations of reality.

Key words: conversation, interaction, matching task, metacommunication

Matching task: a joint activity

The project was aimed at exploring cognitive tools, linguistic techniques and discursive devices spontaneously formed by subjects who participated in an experimental study of interaction in dyads: two subjects accomplished experimental task with an opaque screen preventing them from seeing each other's face and workspace. One of the subjects (Director) explained her partner (Matcher) how to assemble Lego blocks into a model to match the prototype or, in another series, how to arrange photographs of cloudy sky on the table, in order to match the source layout visually available to Director only.

Lack of visual access to the partner's field of operation emphasizes the role of coordination, or alignment, of perspectives between the partners. Participants' construal of each other's view can only rely on spoken discourse, without visual feedback concerning partner's actions and understanding. Accordingly, mechanisms of reference, anaphora, conceptual choice, and deictic anchorage of utterances function in somewhat special ways under these conditions compared to face-to-face conversation in a shared space where pointing gestures and visual monitoring of each other's activity are available.

Matching task paradigm has already been used for the study of communicative interaction since Piagetian studies of referential communication dating back to 1920s, cf. more recent developments in

Wilkes-Gibbs, Clark (1992); Metzing, Brennan (2003); Clark, Krych (2004), and also a concise review in Schober (2006). In our version, in both Lego and photograph tasks, each partner acted as Director and Matcher consecutively. Pairs of subjects comprised normal adults, young children and adult schizophrenic patients, in homogeneous as well as mixed pairs composed of normal adults interacting with schizophrenic adults, and normal adults interacting with children, totalling 200 episodes of interaction, videotaped and transcribed to form a multimedia corpus. The rationale behind involving schizophrenic patients is the claim that schizophrenic patients suffer from a specific impairment of the ability to deduce the mental state of communicative partner, though their cognitive abilities remain unaffected, see Brune (2005) for review of literature. We looked for the features of communicative interaction typical for schizophrenics' less efficient performance in experimental tasks compared to normal controls. We also studied communicative strategies used by more cognitively and communicatively efficient partner in mixed pairs to compensate their partner's deficiency.

Since matching task involves partners' partial access to reality and critically depends on cooperation, it cannot be reduced to partners' individual actions. We consider this to be a case of distributed cognition, as conceived by E.Hutchins (1995). It is essentially public because the control of activity is performed by means of spoken discourse.

Dialogue and the management of activity: meta-level phenomena

During the interaction, partners' contributions to dialogue organize the activity of task accomplishment and also work for internal organization of the dialog itself, providing for turn-taking, display of understanding, monitoring partner's understanding, etc. Dialogue work is particularly evident in the phenomenon of repair (including self-repair, other-repair, and repair initiation). As conceptualised within conversation analytical paradigm, repair demonstrates participants' concern with mutually certified understanding: the absence of repair after the next turn means by default that partner's understanding of the first turn was accepted by the speaker, see, e.g., Schegloff (1992) and elaboration of this logic in the notions of *presentation phase* and *acceptance phase* by H.Clark (1996). Self-repair and related disfluencies can be psycholinguistically explained by internal monitoring of one's own speech production; for classification of self-repair in a corpus of task-oriented dialogues, see studies by Levelt, starting from Levelt (1983).

Studies by Clark and associates have shown that what is going on in matching task dialogues mostly fits within identification and localization phases performed consecutively for all the elements. We have studied

optional verification sequences, as well as some other types of contributions like in this example (quoted in translation from Russian):

M[atcher]: we should win
 D[irector]: wow!
 >M: let's start with the simple ones.
 D: if they were labeled "this is a simple one"... (LAUGHS)
 M: those which have blue spots
 D: you've got four pics with blue spots
 >M: ok. well, here I provisionally set apart some four
 D: four

In the marked lines, utterances have to do with general management of activity, whereas local management can be exemplified with signalling or requesting about readiness to go to a next step. Other meta-level activities considered in this study include: explicit evaluation of the state of joint project, of one's own state or partner's state; retrospective references and accounts; meta-communicative utterances for conversation management; backchannel response: "continuers"; display utterances: collaborative completion, echoic quotation; meta-linguistic utterances; meta-communicative utterances having to do with rapport, etiquette, joking. If present in a participant's speech, these types of contributions to dialogue can be regarded as features of interactivity, involvement, and concern with intersubjectivity.

Results: norms and pathology

Normal participants compared against schizophrenic patients show significantly more relevant utterances. This correlates with greater flexibility in choosing strategies of joint activity. Thus, unlike in schizophrenia, the normal Matchers readily switch from matching to describing, when Director has difficulties in explaining. Some types of utterances better correspond to Director's or Matcher's role, cf., for instance, general activity management seems to correspond to Director's endeavour.

Table 1. General activity management utterances, present in episodes, per cent.

Pathology		Norms	
Directors	Matchers	Directors	Matchers
44 %	22 %	86 %	71 %

Utterances that display understanding and acceptance, particularly, echoic repetition of partner's word(s) and collaborative completion of partner's utterances, also differ in norms and pathology, leading to a conclusion that for some reason cognitively non-impaired schizophrenic patients with fluent speech show less interest in actively updating a common ground in dialogue.

Table 2. Display utterances (collaborative completion, echoic response), average per episode.

Pathology		Norms	
Directors	Matchers	Directors	Matchers
1.0	1.6	2.5	4.5

Pathology demonstrates a deficit of explicit planning and coordination of joint activities, reduced or lacking verifications, and a lack of metalinguistic utterances. Explicit metacommunicative utterances, although differently across their types, are also less numerous in pathology. In mixed pairs, metacommunicative utterances by the normal partner seem to be a means to enhance the dyad's efficiency.

Acknowledgements

Grants from the Russian Ministry of Education and Science (# 16.740.11.0113, 02.740.11.0369)

References

- Brune, M. 2005. "Theory of mind" in schizophrenia: A review of literature. *Schizophrenia Bulletin* 31, 1, 21-42.
- Clark, H.H. 1996. *Using Language*. Cambridge: Cambridge University Press.
- Clark, H.H., Krych, M.A. 2004. Speaking while monitoring addressees for understanding. *Journal of Memory and Language* 50, 62-81.
- Hutchins, E. 1995. *Cognition in the Wild*. Cambridge (MA): MIT
- Levelt, W. 1983. Monitoring and self-repair in speech. *Cognition* 14, 41-104.
- Metzing, C., Brennan, S.E. 2003. When conceptual pacts are broken: partner-specific effects in the comprehension of referring expressions. *Journal of Memory and Language* 49, 201-213.
- Schegloff, E. 1992. Repair after next turn: The last structurally provided defence of intersubjectivity in conversation. *American Journal of Sociology* 97, 5, 1295-1345.
- Schober, M.F. 2006. Dialogue and Interaction. In Brown K., (Editor-in-Chief) *Encyclopedia of Language & Linguistics*, Second Edition, volume 3, 564-571. Oxford: Elsevier.
- Wilkes-Gibbs, D., Clark H.H. 1992. Coordinating beliefs in conversation. *Journal of Memory and Language* 31, 183-194.

Perceptual level of intonation in whispered voice

Géraldine Vercherand

CLILLAC - ARP, UFR Linguistique, Université Paris Diderot, France

Abstract

In this paper, I propose to analyze the prosodic realization of two modalities (declarative and interrogative) through perception study of whispered voice in French. I present results of two modality identification tasks. Results of identification test based on natural stimuli show a good identification rate of the modality. The values of the duration of final syllable seem to be the main acoustic cue that is relevant in the identification of modality. I present also the results of modality identification task based on re-synthesized stimuli allowing to control the role played by duration parameter. Results of this second experiment show modality identification depends on perceptual level of the duration.

Key words : whispered speech, intonation, perception, acoustic parameters.

Introduction

We know that whisper implies no vibration of the vocal folds, provoking thus the absence of fundamental frequency. However intonation continues to be recognized in whisper.

Most of the works on whisper concern its physiological and acoustical aspects (glottal characteristics (Tartter 1989) laryngeal movements (Bonnot 1987) and consonant closure duration (Higashikawa 1996), where generally the exploited material for the analysis consists of isolated words or non-words (Higashikawa 1996). Very few studies deal with the prosodic realization in whisper (Meyer-Eppler 1961) or the differences between voiced speech and whisper in larger items than words such as sentences (Faraco 1983 for French).

My study tries to check results obtained on a corpus more controlled (than Faraco 1983) and to measure the importance of acoustic parameters. So for this study, I will start with 2 hypotheses:

- a) each parameter (values of formants, duration, speech rate and intensity) contributes to the detection of modality ;
- b) rhythmic unit is important to the realization of acoustic parameters.

Natural stimuli

The corpus

The French corpus consists in 11 phrases pronounced as declarative and interrogative. The length was incremented from 2 syllables to 12 syllables, 1 syllable incremented each time.

Table 1 : Corpus (Combination of an element per column).

Elle (1,2) Anna (3,4,5,6) Annabelle (7,8,9) Annabella (10,11)	mange	Ø (1) bien (2,3) une noix (4) un melon (5) des ananas (6,7) une noix du Ghana (8) un melon du Ghana (9,10) des ananas du Ghana (11)
--	-------	--

Perception test

This perception test based on natural stimuli aims a) to determine if modality is recognized in whispered speech, b) to see if the sentence length (number of syllables) has an effect on the rate of recognition of modality and c) to understand how modality is realized and to determine which acoustics phenomena are decisive.

Perception test results

The modality is fairly recognized in whispered speech. We had 78% response as result. The ANOVA test shows statistically significant effect by length on identification rate ($F(10,198)=5,421$, $p<0.0001$).

Confirming my hypothesis, the perception rate correlates with the phrasal length (Interrogative : 2 to 5 syllables 67,5%, 6 to 8 syllables 79,5%, 9 to 12 syllables 89,2% ; Declarative : 2 to 5 syllables 68,5%, 6 to 8 syllables 79%, 9 to 12 syllables 82,75%).

Analyze

A total of 6 items, 3 interrogative and 3 declarative were less recognized by all listeners.

- *declarative sentences*. The value of speech rate and final duration are different from the average expected. Duration seems to be the most important parameter.

- *interrogative sentences*. Several parameters seem to be involved, as too slow speech rate, duration of the last syllables too long and a dynamic intensity too high.

Resynthesized stimuli

Perception study based on natural stimuli shows that the value of a parameter seems to compensate the other and vice versa it could alter interpretation. So I choose to vary the duration of the last syllable in order to verify its influence on the identification scores. I hypothesize that increasing the duration of the last syllable compensates the absence of other acoustic parameters.

The corpus

I select three declarative sentences of test corpus based on natural stimuli. These sentences are representative of the number of syllables (3 syllables, 7 syllables, 12 syllables).

I built a continuum of time on the last syllable. The continuum is composed by 5 steps. The synthesized stimuli were generated from the third stimulus of the continuum (Stimulus 1 -80ms, S2 -40ms, S4 +40ms, S5 +80ms). The corpus is composed of 5 stimuli per each sentence. Each stimulus appears three times in the test, representing a total of 45 stimuli.

Perception test

The task of the test is to identify the modality. The listeners are asked to choose among three possibilities: “declarative”, “interrogative” and “I do not know”. 12 French listeners participate. The listeners take first a pre-test to familiarize themselves with the type of stimuli. The stimuli are presented randomly.

Perception test results

Table 3 : Identification scores of modality with synthesized stimuli - in row stimuli, in column responses (%).

Sentences	Simuli	Declarative	Interrogative	I do not know
Short	1	61,11	36,11	2,78
	2	41,67	58,33	0
	3	16,67	83,33	0
	4	8,33	91,67	0
	5	19,44	77,78	2,78
Medium	1	88,89	11,11	0
	2	94,44	2,78	2,78
	3	86,11	8,33	5,56
	4	80,56	13,89	5,56
	5	69,44	27,78	2,78
Long	1	41,67	50	8,33
	2	44,44	55,56	0
	3	25	69,44	5,56
	4	27,78	61,11	11,11
	5	25	66,67	8,33

More duration of the last syllable increases, more listeners identify the sentence as interrogative. An ANOVA test shows a significant effect of treatment duration ($F(4,15)=1,538E15$, $p<0,0001$), a significant effect of sentence type ($F(2,15)=1,551E16$, $p<0,0001$) and an interaction between the two independent factors ($F(8,15)=3,143E15$, $p<0,0001$).

Thus increased duration of the last syllable influence significantly the recognition rate. The level of duration to identify modality seems to depend on the length of the sentence. The length of the sentence influences the recognition rate and interacts with changes in duration.

Discussion and conclusion

Through this study on whispered speech, it shows that each parameter (duration, speech rate, intensity and formant) involved in the perception of the two modalities patterns, the declarative and interrogative. The length of the sentences parameter is important. It seems that the rhythmic factor plays an important role, since a number of syllables is necessary to allow the parameters of intensity and duration to differentiate sufficiently and to allow perception of different modalities. In addition, this study shows that the prominence of an acoustic parameter compensates for non-prominence of others in the detection of prosodic phenomena studied. However, the required level seems to depend on the length of the sentence. Further studies are underway to explore this question and to widen of perspectives to other modalities.

References

- Bonnot J-F.P. 1987. Electromyographical Correlates of Shouted and Whispered Voice. Actes du 11ème Congrès International des Sciences Phonétiques, volume 4, 20-23, Tallinn, Estonie, 1987.
- Faraco M. 1983. Comparaison des intonations affirmatives et interrogatives en voix normale et voix chuchotée, Thèse de 3ème cycle, Université d'Aix en Provence.
- Higashikawa, & al. 1996. Perceived Pitch of whispered voice—Relationship with formant frequencies: a preliminary study. *Journal of Voice* 10.2: 155-158.
- Meyer-Eppler, W. 1957. Realization of prosodic features in whispered speech. *JASA* 29, 104-106.
- Tartter V.C.1989. What's in whisper? *JASA* 86 (5), 1678-1683.

Index of authors

Ankerstein, C.A.	11	Kireev, M.	59
Bahadır, G.	15, 31	Korotkov, A.	59
Bakakou-Orphanou, Aik.	39	Madureira, S.	51, 115
Bánréti, Z.	19, 23	Magoula, E.	43, 111
Bardiaux, A.	47	Martin, Ph.	95
Blochowiak, J.	31	Medvedev, Sv.	59
Bonitz, P.-K.	35	Memetova, K.	59
Botinis, A.	39, 43	Mendousse, K.	99
Camargo, Z.A.	51	Mészáros, E.	19, 23
Chaida, A.	39, 43, 143	Mykhaylyk, R.	103
Chanethom, V.	55	Nikolaenkova, O.	107
Chasan, O.	111	Nirgianaki, E.	111
Chernigovskaya, T.	59	Omana, M.	131
Chernigovskaya, T.	147	Őrley, Z.	23
Collado, J.	131	Polinsky, M.	15
de Camargo, Z.A.	115	Rusilo, L.C.	51, 115
de Mareüil, Ph.B.	47	Sang, Y.	119
Di Russo, D.	63	Sloos, M.	123
Dubasova, A.	67	Soroli, E.	127
Fecher, N.	71	Thompson, E.	131
Galantomos, I.	75	Todd, R.	135
Gor, K.	59	Toplu, A.B.	27
Guillamón, P.M.	79	Torrens, V.	139
Hofferberth, N.J.	83	Tripolitou, M.	143
Hoffmann, I.	23	Utekhin, I.	147
Holler, A.	35	Vercherand, G.	151
Karpava, S.	87	Xu, Y.	1
Kataeva, G.	59	Yousuf-Little, A.	131
Kim, M.	91	Zeyrek, D.	27



ExLing 2011

Ebook ISSN: 2529-1092

Ebook ISBN: 978-960-466-158-9