



International Society of Experimental Linguistics

## ExLing 2019

### Proceedings of 10<sup>th</sup> International Conference of Experimental Linguistics

25-27 September 2019  
Lisbon, Portugal

Edited by Antonis Botinis







International Society of Experimental Linguistics

## **ExLing 2019**

Proceedings of 10<sup>th</sup> International Conference on  
Experimental Linguistics  
25-27 September 2019, Lisbon, Portugal

Edited by Antonis Botinis



ExLing 2019  
Proceedings of 10<sup>th</sup> International Conference of Experimental Linguistics

Athens, Greece  
ISSN: 2529-1092  
ISBN:  
Published by ExLing Society  
Copyright © 2019 ExLing Society and authors

## **Foreword**

---

This volume includes the proceedings of ExLing 2019, the 10<sup>th</sup> International Conference of Experimental Linguistics in Lisbon, Portugal, 25-27 September 2019. The first ExLing Conference, at the time an International Speech Communication Association (ISCA) ExLing Workshop, was organised in Athens, Greece, in 2006, and has been regularly held thereafter in different places.

Since 2015, the ExLing Conference is held annually and since 2019, on the occasion of the set up of the International Society of Experimental Linguistics, it is called the International Conference of Experimental Linguistics and simply ExLing Conference.

In accordance with the spirit of the ExLing conferences, new and established researchers came to Lisbon to discuss developments in linguistic research and experimental methodologies. This time, we included in the programme group discussions too, so that ExLing participants had the opportunity of talking with each other in a relaxed and constructive context.

We are happy to see that our initial attempt is more and more becoming an established forum for new generations of linguists. As in the previous conferences, our colleagues have come from different parts of the world and we hope they have had a rewarding exchange of scientific achievements and expertise. This is indeed the core of the ExLing Conferences, which promotes new ideas and methodologies in an international milieu.

We would like to thank all ExLing 2019 participants as well as the Centro de Linguística at the University of Lisbon. We also thank our keynote speakers, Bart Geurts, Jonathan Harrington, Marc Swerts and colleagues from the International Advisory Committee as well as Conference assistants for their contribution to the successful outcome of the Conference.

Antonis Botinis



## Contents

---

### Keynote papers

<i>What's wrong with Gricean pragmatics</i> .....	1
Bart Geurts .....	1
<i>Linguistic adaptation</i> .....	9
Marc Swerts.....	9
<i>L2/L3 acquisition of the English indefinite article</i> .....	13
<sup>1</sup> Asma Al-hawi, Sviatlana Karpava <sup>2</sup> .....	13
<i>L1 and L2 skills in Specific Reading Disorder</i> .....	17
Georgia Andreou, Vassiliki Tsela, Maria Liakou.....	17
<i>Discourse in Mild Cognitive Impairment</i> .....	21
Malin Antonsson .....	21
<i>Phonosemantic shape-sound correlates</i> .....	25
Elena Besedina .....	25
<i>Tonology and sentence structure in Greek</i> .....	29
Antonis Botinis <sup>1</sup> , Athina Kontostavlaki <sup>1</sup> , Olga Nikolaenkova <sup>2</sup> , Charalambos Themistocleous <sup>3,4</sup> .....	29
<i>Grammatical mood and ambiguity aversion</i> .....	33
Sacha Bourgeois-Gironde <sup>1</sup> , Alda Mari <sup>1</sup> , David Nicolas <sup>1</sup> , David Blunier <sup>2</sup> .....	33
<i>Affective analysis of customer service calls</i> .....	37
Vera Cabarrão <sup>1,2</sup> , Mariana Julião <sup>1,3</sup> , Rubén Solera-Ureña <sup>1</sup> , Helena Moniz <sup>1,2,4</sup> , Fernando Batista <sup>1,5</sup> , Isabel Trancoso <sup>1,3</sup> , Ana Isabel Mata <sup>2</sup> .....	37
<i>Interactions of voice quality settings</i> .....	41
Zuleica Camargo <sup>1</sup> , Albert Rilliard <sup>2</sup> .....	41
<i>On the argumenthood of optional PPs with Italian motion verbs</i> .....	45
Giulia Cappelli <sup>1</sup> , Pier Marco Bertinetto <sup>1</sup> , Alessandro Lenci <sup>2</sup> .....	45
<i>Russian homophones and visual word recognition</i> .....	49
Daria Chernova.....	49
<i>Image effects and long distance dependencies</i> .....	53
Leticia M.S. Corrêa <sup>1</sup> , Marina R.A. Augusto <sup>2</sup> .....	53
<i>On the processing of object relative clauses</i> .....	57
Leticia M.S. Corrêa <sup>1</sup> , Erica dos S. Rodrigues <sup>1</sup> , René Forster <sup>2</sup> .....	57
<i>Automatic administration of semantic verbal fluency tests for Portuguese</i> .....	61
João Costa, Ana Pereira, Liliana Ferreira.....	61
<i>Auditory priming of pseudo-suffixed words</i> .....	65
Ava Creemers, Nattanun Chanchaochai, David Embick .....	65
<i>Neurodegeneration and the Gothenburg MCI study</i> .....	69
Marie Eckerström, Anders Wallin.....	69
<i>Corpora and language variation in Greek</i> .....	73
Fatima Eloeva <sup>1</sup> , Maxim Kisilier <sup>2</sup> , Olga Nikolaenkova <sup>3</sup> .....	73
<i>Stream analysis for detecting stuttering episodes</i> .....	77

Fabio Fassetti <sup>1</sup> , Iliaria Fassetti <sup>2</sup> , Simona Nisticò <sup>1</sup> .....	77
<i>Vocal and facial expressions and meaning effects in speech expressivity</i> .....	81
Mario Fontes, Sandra Madureira .....	81
<i>Reading and Mild Cognitive Impairment</i> .....	85
Kristina Lundholm Fors <sup>1</sup> , Malin Antonsson <sup>1</sup> , Dimitrios Kokkinakis <sup>1,2</sup> , Kathleen C. Fraser <sup>3</sup> .....	85
<i>Italian vowel production by Arabic speakers</i> .....	89
Manuela Frontera <sup>1</sup> , Emanuela Paone <sup>2</sup> .....	89
<i>Garden-path sentences and executive functions in normal aging</i> .....	93
Larissa Rangel Ferrari <sup>1</sup> , Erica dos Santos Rodrigues <sup>1</sup> , Daniel Correa Mograbi <sup>2</sup> .....	93
<i>Iconicity in grammatical variation</i> .....	97
Takashi Fujiwara <sup>1</sup> , Fuminori Nakamura <sup>2</sup> , Daisuke Suzuki <sup>1</sup> .....	97
<i>Lexical deterioration in Alzheimer's disease</i> .....	101
Frédérique Gayraud <sup>1</sup> , Isabelle Simoes Loureiro <sup>2</sup> , Camille Frouin <sup>1</sup> , Cynthia Collette <sup>2</sup> , Laurent Lefebvre <sup>2</sup> , Melissa Barkat-Defradas <sup>3</sup> .....	101
<i>Lexical variation in Belarusian Russian</i> .....	105
Olga Goritskaya, Alexandra Chudar .....	105
<i>Vocabulary internationalisms and start-ups</i> .....	109
Agnieszka Grażul-Luft .....	109
<i>Experimental linguistics in aphasia and child language</i> .....	113
Maria Irma Hadler Coudry, Isabella de Cássia Netto Moutinho .....	113
<i>Lexical diversity and Mild Cognitive Impairment</i> .....	117
Sofie Johansson <sup>1</sup> , Kristina Lundholm Fors <sup>1</sup> , Malin Antonsson <sup>1</sup> , Dimitrios Kokkinakis <sup>1,2</sup> .....	117
<i>Linguistic Intraoperative Protocol</i> .....	121
Luis M.T. Jesus <sup>1,2</sup> , Joana Alves <sup>1</sup> , Mafalda Cardoso <sup>1</sup> , Mariana Morgado <sup>1</sup> .....	121
<i>Animacy distinctions and semantic dementia</i> .....	125
Anna Kaglik <sup>1</sup> , Clara Sanches <sup>1</sup> , Elsa Mhanna <sup>1</sup> , Raffaella Migliaccio <sup>1,2</sup> , Marc Teichmann <sup>1,2</sup> , Antoni Valero-Cabré <sup>1</sup> .....	125
<i>Linguistic and extra-linguistic triggers in intrasentential codeswitches</i> .....	129
Yuliya Leshchenko, Tatyana Ostapenko .....	129
<i>Automatic detection of language impairment</i> .....	133
Hali Lindsay <sup>1</sup> , Johannes Tröger, Nicklas Linz <sup>1</sup> , Jan Alexandersson <sup>1</sup> , Johannes Prudlo <sup>2</sup> .....	133
<i>The use of pronominal forms in complement clauses in European and Brazilian Portuguese</i> .....	137
Paula Luegi <sup>1</sup> , Marcus Maia <sup>2</sup> , Armanda Costa <sup>1</sup> .....	137
<i>Deep learning and intonation in Text to Speech systems</i> .....	141
Philippe Martin .....	141
<i>Transcranial stimulation in semantic dementia</i> .....	145

Elsa Mhanna <sup>1</sup> , Clara Sanches <sup>1</sup> , Antoni Valero-Cabr� <sup>1</sup> , Marc Teichmann <sup>1,2</sup> .....	145
<i>Semantic abilities and narration in aphasia in Greek</i> .....	149
Fani Nasika.....	149
<i>Narrative development and gestures in different genres</i> .....	153
G�khan �zkayin.....	153
<i>Simulating alphabet recitation under thalamic lesions</i> .....	157
Martin D. Pham <sup>1</sup> , Terrence C. Stewart <sup>2</sup> , Randy Harris <sup>3</sup> .....	157
<i>Emotional prosody perception in Italian as a second language</i> .....	161
Emanuela Paone <sup>1</sup> , Manuela Frontera <sup>2</sup> .....	161
<i>Experimental customisation of the Versioning Machine</i> .....	165
Elsa Pereira <sup>1</sup> , Roman Bleier <sup>2</sup> .....	165
<i>Camilian lexical substitutions</i> .....	169
Carlota Pimenta.....	169
<i>Spatial anaphora in Hungarian</i> .....	173
Gy�rgy R�kosi, Enik� T�th.....	173
<i>The acoustic characteristics of the Baraba-Tatar and Shor consonant ‘r’</i> .....	177
Tatiana Ryzhikova <sup>1,2</sup> , Nikolay Urtegeshev <sup>1,2</sup> .....	177
<i>Processing the meaning of English articles by Arab learners</i> .....	181
Mona Sabir, Alaa Melebari.....	181
<i>Neurodegenerative language impairments and transcranial stimulation</i> .....	185
Clara Sanches <sup>1</sup> , Juliette Godard <sup>1</sup> , Marc Teichmann <sup>1,2</sup> , Antoni Valero-Cabr� <sup>1</sup> .....	185
<i>Lexicalized pauses in Italian</i> .....	189
Loredana Schettino, Violetta Cataldo.....	189
<i>The natural environment for the experimental study of phonosemantics</i> .....	193
Elena Shamina.....	193
<i>Analysing L2 Swedish word-final stops</i> .....	197
Maryann Tan <sup>1</sup> , Xin Xie <sup>2</sup> , T. Florian Jaeger <sup>2</sup> .....	197
<i>Speech and Mild Cognitive Impairment detection</i> .....	201
Charalambos Themistocleous <sup>1,2</sup> , Dimitrios Kokkinakis <sup>2,3</sup> .....	201
<i>Hungarian demonstratives in contrastive contexts</i> .....	205
Enik� T�th.....	205
<i>Lexical selection and gender agreement processing</i> .....	209
Masayuki Yamada, Maria do Carmo Louren�o-Gomes.....	209
<i>Interpreting Mandarin positive polar questions</i> .....	213
Chenjje Yuan, Peng Li.....	213



# What's wrong with Gricean pragmatics

Bart Geurts

Department of Philosophy, Radboud University, Netherlands

DOI:

## Abstract

The view that human communication is essentially a matter of sharing mental states, especially communicative intentions, has been immensely influential in pragmatics and beyond. Drawing together and elaborating various lines of criticism, I argue that this influence has been mostly harmful; in particular, it has misdirected research on the evolution and development of language and communication.

Key words: Gricean pragmatics, mentalism, communicative intentions

## Introduction

It is widely held that communication is a form of information exchange. Pragmatic theories usually embrace a very specific version of this doctrine. To begin with, “information” is restricted to “what’s in the speaker’s mind”; in this sense, the field of pragmatics is predominantly *mentalist*. Moreover, it is broadly accepted that, of all the kinds of things that may be in the speaker’s mind, one is the payload of linguistic communication, namely, communicative intentions; in this sense, pragmatic theories are mostly *Gricean*, for the notion of communicative intention (not the name) goes back to Grice’s (1957) theory of speaker meaning. Another label for this view is “intentionalist”, which I will use interchangeably with “Gricean”.

Since 1957, there has developed a cottage industry of revising and rerevising Grice’s original account, but fortunately we need not concern ourselves with that. Here it will suffice to say that a communicative intention is meant to be recognised as such by the hearer, and is aimed at causing some change in the hearer, typically a change in his belief set. For example, if Alice tells Bruce, “The goose is done”, intending him to come to believe, due to his recognising her intention, that the goose is done, then Alice communicatively intends Bruce to believe that the goose is done.

The intentionalist mantra is that communication is a matter of expressing communicative intentions, on the part of the speaker, and recognising such intentions for what they are, on the hearer’s part. This view is widespread and has been shaping debates and research programmes not only in linguistics and philosophy, but also in the study of language development and language evolution. If there is a dominant view on human communication, this is it.

The following quote illustrates the popularity of intentionalism (as well as the hubris that popularity is liable to breed):

The whole of modern pragmatics is predicated on this assumption [that intentionalism is right], and its findings are arguments in favour of it. Of course, this does not make the assumption right, but those who deny it, are, in effect, implying that pragmatics as currently pursued is a discipline without an object, somewhat like the study of humours in ancient medicine. (Origgs and Sperber 2000: 156)

If Origgs and Sperber are to be believed, there must be quite a number of charlatans out there. I'm one of them.

With all due respect to Origgs and Sperber, I find precious little evidence to support intentionalism. I'm willing to grant that many, though not all, utterances license inferences about the speaker's mental state; but that doesn't begin to show that it is their chief purpose to express the speaker's mental state. After all, even if utterances typically convey information about the speaker's gender, that is obviously not what they are for. Moreover, contrary to what Origgs and Sperber claim, there are perfectly viable alternatives to the Gricean approach (e.g. Brandom 1994, Geurts 2019).

The objective of this paper is to draw together and elaborate various lines of criticism that have been levelled against mentalist pragmatics, and especially against its Gricean varieties.

### **A dubious dichotomy**

It is agreed by friend and foe alike that the speaker's mental state is not always relevant to the successful performance of a speech act.

(1) I hereby declare the cricket season open.

For an utterance of (1) to be felicitous, there must be an accepted convention to the effect that the cricket season is to be opened at a certain time, by a person who is licensed to do so, using a formula along the lines of (1). If all these conditions are met, an utterance of (1) will be successful, and the cricket season will have been opened. The speaker's mental state is irrelevant, and that includes his communicative intentions, if there be any. There is a consensus on this, and advocates of Gricean pragmatics have solved the problem by gerrymandering: they stipulate a dichotomy between speech acts proper, to which their account applies, and "conventional" speech acts, for which communicative intentions are irrelevant.

On the face of it, this tactic may seem sound, because examples like (1) appear to be special, and a theory shouldn't be blamed for failing to deal with deviant cases. However, on reflection, the dichotomy between proper and conventional speech acts is dubious. Consider permission giving: I grant you permission to use my fountain pen by saying: "You can use my fountain pen." Arguably, it is a convention in English-speaking communities that

permission to do so-and-so may be granted by a speaker provided certain conditions are fulfilled; notably, in the case at hand, I must own the fountain pen in question. If these conditions are met, you will have obtained permission to use my fountain pen, no matter what my mental state may be. This case suggests that giving permission is like opening the cricket season in all relevant respects, and similar cases can be constructed with promises or questions, for example. This is not to imply that it is established to everybody's satisfaction that all these speech act types are conventional; but it is to say that, at the very least, the dichotomy between conventional and non-conventional speech acts is a moot one. It is up to the Griceans to prove that this distinction is not merely an artefact of their view on communication.

### **Developmental issues**

An oft-voiced worry about intentionalism is that it raises the bar for successful communication so high that it becomes impossible to account for the basic fact that children learn to use language for communication well before their school years. If understanding a simple statement like "It's cold" requires an understanding of the speaker's intentions and beliefs, children must be able to attribute these mental states before they can understand statements. But, on the face of it at least, attributing intentions and beliefs is hard, and attributing communicative intentions must be even harder. There is no independent evidence to show that three-year olds have these skills, and such experimental evidence as is currently available suggests that they don't. Therefore, intentionalism seems to entail that preschoolers must be bad at linguistic communication, which is not the case.

This argument hinges on two premisses that may be, and have been, questioned: that attributing intentions and beliefs is hard, and that there is good reason to doubt that three-year olds can attribute these mental states. We'll get to these points presently, but first let's develop the argument in a different direction, starting from the uncontroversial fact that, as things currently stand, there are no well-articulated theoretical models of how children might learn to attribute intentions and beliefs before they start communicating linguistically. That being so, it is tempting to stipulate that mental-state attribution doesn't have to be learned in the first place: it's in the genes. This flight into nativism amounts to passing the buck from development to evolution, because now the question is how our ancestors could have evolved the skill of attributing beliefs and intentions before they were capable of communicating linguistically. And then there is also the matter of explaining how this ability got encoded in our genetic material. Again there are no well-articulated models of how all of this might have happened, which is not surprising, since the phylogenetic problem must be at least as hard as its ontogenetic counterpart.

There are ideas and even theories about how the attribution of beliefs and intentions develops early on in life once children have acquired basic communicative skills (e.g. Jary 2010, van Cleave and Gauker 2010), and some of these ideas and theories may be retooled to develop an evolutionary account of mental-state attribution (Geurts, in preparation). But clearly this line of thinking reverses the direction of explanation required by mentalist pragmatics, proceeding as it does from communication to attribution instead of the other way round. In the absence of theories that proceed in the opposite direction, this is an argument against mentalism in general, and Gricean pragmatics in particular.

### **Mental-state attribution is no easy matter**

Communicative intentions are higher-order mental states: they are intentions targeted at other mental states, which themselves may be targeted at other mental states. Moreover, communicative intentions are self-referential: the speaker intends the hearer to recognise that very intention. Hence, communicative intentions are complex structures, which is a good reason for supposing that the ability to attribute communicative intentions is hard to acquire, be it ontogenetically or phylogenetically.

But there is another reason as well. Beliefs and intentions are highly abstract in the sense that, generally speaking, they are not readily inferable from their owners' behaviours. Does Alice think that 43 is an irrational number? You may observe her actions for the rest of her life (or your own), and never find out. Ditto for intentions, even if they are at a lesser remove from overt action than beliefs. In this respect, there is a stark contrast with pain, with its characteristic external causes and manifestations. Moreover, pain is associated with salient and distinctive experiential states, which are lacking in beliefs and intentions, as well. Even if the experience of pain is impossible to verbalise, we all know what it is like to have pain; but what is it like to think that 43 is an irrational number? To my mind, the best and most likely response to this question would be a helpless shrug. Given that pain has a distinctive and salient phenomenology, is reliably caused by certain external events, and manifests itself externally in stereotypical ways, it is not too hard to imagine how a child might learn to attribute pain. Beliefs and intentions are quite a different matter.

So there are at least two major challenges for developmental and evolutionary theories of Gricean communication: the abstractness of beliefs and intentions and the complexity of communicative intentions. Recent attempts at simplifying communicative intentions or decomposing them into simpler units that might evolve or be acquired separately leave the first challenge unaddressed (Moore 2017, Sterelny 2018).

The foregoing argument is theoretical; it doesn't rule out the possibility that, as a matter of empirical fact, attributing beliefs and intentions is not as hard as might be expected for theoretical reasons. As it turns out, some authors support the view that the experimental record shows just that, and that infants have the conceptual resources for dealing with communicative intentions, after all. However, the experimental measures that can be used in studies with infants and toddlers are, of necessity, very indirect, or "implicit", as it is sometimes called. The jump from, e.g., looking-time patterns to belief attribution is quite a big one, and not everybody is prepared to jump that far. Moreover, the quality of the data on which this line of argument hinges is dubious, and that's putting it quite mildly. As discussed by Rakoczy and Behne (2019), most of the experimental measures on which the evidence is based have failed to replicate, and to the extent that they do replicate, their validity is in doubt. "In light of the existing evidence, we do not know whether there is indeed such a thing as implicit [attribution of beliefs and intentions] before age 4." (Rakoczy and Behne 2019: 94)

Summing up: while there are good reasons for supposing that attributing beliefs and intentions is hard and that attributing communicative intentions is even harder, the data that have been claimed to undermine these claims are controversial, as is their interpretation. Hence, the prospects for an account of the ontogenesis and/or phylogenesis of Gricean communication are bleak.

### **Pragmatics is not mind reading**

At the same time that Gricean pragmatics was coming into the ascendant, researchers in psychology, primatology, and the philosophy of mind began discussing mental-state attribution in terms of a variety of new epithets, including "mind reading", "theory of mind", and "mentalising" (I find these labels about equally bad, but will stick with the first). The experimental workhorse of this development was the false-belief task, which became the yardstick for mind-reading ability. The connection between mind reading and mentalist pragmatics was impossible to overlook, and soon the two were regularly conflated: "This paper defends the broadly Gricean view that pragmatic interpretation is ultimately an exercise in mind-reading, involving the inferential attribution of intentions." (Sperber and Wilson 2002: 3)

It seems to have gone unnoticed that these developments gave mentalist pragmatics a radically new turn. The leading idea of mentalist pragmatics was and is that communication is a matter of expressing and recognising mental states, by speaker and hearer, respectively. This was welded to the notion that mind reading does the heavy lifting in pragmatic interpretation. These two ideas are patently distinct, and it is unfortunate that they got mixed up, because, whereas the first idea is not completely off the mark, the second is. As I said before, I consider it plausible that many utterances

license inferences about speakers' mental states. This is a mentalist position of sorts, though considerably weaker than the industrial-strength mentalism under discussion. But the notion that most or all pragmatic reasoning is mind reading is supported by neither evidence nor reason.

It's a harsh verdict, I know, but I'll try to back it up by way of a handful of miniature case studies, each of which represents a different kind of puzzle that any pragmatic theory should be able to solve. Here's case number one:

- (2) The ship's engine started to splutter, and the captain decided to head for the nearest port.

Although the word "port" is ambiguous, in (2) there is a strong preference for reading it as "harbour" rather than "fortified red wine", for example. Why? Because this yields the most coherent interpretation for the sentence. This doesn't contradict the view that pragmatic interpretation is an exercise in mind reading, but it isn't evidence for that view, either, because coherence is bound to be a constraint on interpretation for *every* pragmatic theory.

Case number two:

- (3) Bruce was angry with Clyde, because he stole his fountain pen.

Who stole whose fountain pen? The answer is obvious, and for the same reason as before: if we take "he" and "his" to refer to Clyde and Bruce, respectively, we obtain a much more coherent interpretation than if we take it to be the other way round, and as in the previous case there is no reason to suppose that mind reading is involved in the process.

Case number three: I tell you a story about a farmer who caught a burglar in his barn, ending my tale thus:

- (4) The farmer took his shotgun and killed the burglar.

You interpret my utterance as implying that the farmer killed the burglar with his shotgun, even though (4) doesn't state this explicitly. How is that possible? This case is a bit trickier than the first two, but it is clear that, again, coherence is the main factor. Presumably, the conjunction "and" indicates that there is some sort of connection between the farmer's actions, and the supposition that the shotgun was used in the killing establishes such a connection in the most obvious way. No mind reading is required.

More cases could be produced at will, but I trust that my point is clear and strong enough: the notion that pragmatic interpretation is mind reading is a non-starter. This is not to imply that mind reading doesn't play a role in pragmatics. Some stories are impossible to make sense of unless we make substantial suppositions about the protagonists' desires, delusions, motives, and so on. Moreover, once derived, an interpretation, may always be capped by concluding that it was so intended by the speaker; but this is a flourish added after the fact, not a key step in the interpretative process. There is no reason to suppose that mind reading is the driving force of pragmatics.

## Idle wheels

A notable feature of the ongoing discourse on communicative intentions is that it tends to focus on *non*-linguistic communication, a trend set by Grice's (1957) landmark paper and continued throughout the subsequent literature. A popular example is holding up an empty glass to signal to your host that a refill would be appreciated (e.g. Scott-Phillips 2015). One wonders how such cases bear on linguistic communication, but that is never properly explained.

Moreover, it is telling that despite the numerous incarnations which the concept of communicative intention has gone through, its explanatory accomplishments have thus far been decidedly modest. Going through the main topics that pragmatics is about (anaphora, definiteness, presupposition, implicature, and so on), I can think of only one which is regularly associated with communicative intentions, viz. speech acts, and even in this part of the pragmatic literature communicative intentions tend to be idle wheels in the theoretical machinery, as we have seen. It should also be noted that in his own work on presupposition and implicature Grice never even mentions communicative intentions (by this or any other name).

## Conclusion

While I do see a place in pragmatics for some measure of mentalism, I don't have high hopes for intentionalism. All things considered, it seems to me that Gricean pragmatics has done much harm and little good. In particular, it has obstructed progress in theorising about the ontogeny and phylogeny of language and communication. But to end on a happy note, I will conclude by considering to what extent mentalism is on the right track.

I accept that in many cases the hearer will be entitled to infer that the speaker intended her utterance to mean such-and-such, that she believes it to be true that such-and-such, and so on. For example, in the case of (4), the hearer is normally entitled to infer that the speaker believes that the farmer killed the burglar with his shotgun. What's more, this inference is normative: given that she uttered (4), the speaker is *supposed* to have this belief. This much mentalism can be granted while maintaining that we get much of our pragmatic business done without attributing mental states to each other.

To make this a bit more concrete, let me briefly sketch an alternative approach, which accommodates a degree of mentalism, but views communication in a non-mentalist way (see Geurts 2019 for details). Instead of trying to penetrate into the innards of speakers' and hearers' minds, this approach starts at the social level, where speech acts are embedded in interactions that are thoroughly normative. A promise, for example, is a way of undertaking an obligation. It may *also* be a way of expressing an intention, and somehow this aspect will have to be accounted for; but it will be derivative of the social-normative function of promising.

Since the undertaking of an obligation is a social act that creates a normative fact, promising need not involve mental-state attribution. However, my promise doesn't count for much if you have reason to suspect that I might not intend to honour my obligation. This is where mental-state attribution comes in: if you are willing to suppose that I am a cooperative speaker, you are entitled to infer that I intend to make good on my promise.

This view accords an important role to mental-state attribution without making it the hub of linguistic communication. Thus it becomes possible to hold that the development of mental-state attribution is contingent on, and must therefore be preceded by, a substantial level of communicative know-how; and the same may be held for the evolution of mental-state attribution.

## Acknowledgements

For comments and discussion I'm indebted to Antonis Botinis, Martina Valković, and Huub Vromen.

## References

- Brandom, R. 1994. *Making it explicit*. Harvard University Press.
- Geurts, B. 2019. Communication as commitment sharing: speech acts, implicatures, common ground. *Theoretical linguistics* 45, 1-30.
- Geurts, B. (in preparation). First saying, then believing: the social roots of folk psychology.
- Geurts, B., Kissine, M., van Tiel, B. 2019. Pragmatic reasoning in autism. In: Morsanyi, K., Byrne, R. (eds.), *Thinking, reasoning and decision making in autism*, 113-134. Routledge.
- Grice, H. P. 1957. Meaning. *Philosophical review* 66: 377-388.<sup>[1]</sup><sub>SEP</sub>
- Jary, M. 2010. Assertion and false-belief attribution. *Pragmatics and cognition* 18, 17-39.
- Moore, R. 2017. Gricean communication and cognitive development. *Philosophical quarterly* 67, 303-326.
- Origg, G., Sperber, D. 2000. Evolution, communication and the proper function of language. In Carruthers, P., Chamberlain, A. (eds.), *Evolution and the human mind*, 140-169.
- Rakoczy, H., Behne, T. 2019. Commitment sharing as crucial step towards a developmentally plausible speech act theory? *Theoretical linguistics* 45, 93-97.
- Scott-Phillips, T. 2015. *Speaking our minds*. Palgrave Macmillan.
- Sperber, D., Wilson, D. 2002. Pragmatics, modularity and mind-reading. *Mind and language* 17, 3-23.
- Sterelny, K. 2018. Language: from how-possibly to how-probably? In Joyce, R. (ed.), *The Routledge handbook of evolution and philosophy*, 120-135.
- van Cleave, M., Gauker, C. 2010. Linguistic practice and false-belief tasks. *Mind and language* 25, 298-328.

# Linguistic adaptation

Marc Swerts

Department of Communication and Cognition, Tilburg University, The Netherlands

DOI:

## Abstract

Conversations are characterised by the fact that dialogue partners adapt their language to specific communicative contexts, where they have been shown to be especially sensitive to characteristics of the addressee. Prior work revealed that this holds for various kinds of linguistic structure, as it affects speakers' pronunciations, lexical choices, and syntactic structures. The current paper presents an overview of related research into gestures, i.e., movements of fingers, hands and arms that speakers produce with speech. We show how such gestures are affected by (1) the type of addressee, (2) the perspective of the addressee and (3) whether a gesture is meaningful for the addressee. We discuss the relevance of these findings for a more general model of spoken interaction.

Key words: Gestures, interactions, adaptation, non-verbal communication.

## Introduction

Speakers adapt their speaking style to various communicative contexts; an important factor is the type of conversation partners with whom they are interacting. For instance, people may use slang words when talking to close friends or family members, but probably feel inclined to a more formal vocabulary when discussing work-related issues with their superiors at their office. Likewise, compared to their default language use, they may use simpler words and less complex syntactic structures, when they are addressing a child or when noticing that their addressee is a non-native speaker who has revealed problems of understanding during prior exchanges. There is much work on how such adaptive behaviour manifests itself, where it has been shown to affect linguistic structures, including pronunciation, lexical choices and syntactic structures (Branigan et al. 2010).

Such adaptive behaviour also appears from speakers' non-verbal features. For instance, it has been shown that speakers differ in their bodily expressions when addressing adults or children. Swerts and Kraemer (2010) compared newsreaders on Dutch public TV broadcasts, who were either reading the news for adults or for children. With children, newsreaders were more clearly signalling the emotional content of a news item through their facial expressions, e.g. clearly distinguishing between news with positive or

more negative content, whereas their expressions were less variable when news with such different emotional connotations was delivered to an adult audience. Conversely, children appreciate a more exaggerated style. House et al. (1999) found that children were more interested in a conversational agent that produced utterances with extreme intonation patterns, when compared to a neutral prosodic style of a model for standard speech synthesis systems.

Thus, features of the addressee may trigger speakers to modulate their nonverbal expressions. The current paper will embark on the gestures that people produce during their spoken interactions, in particular movements of the face, hand and arms (Kendon 2004, McNeill 2005). It is intuitively clear that such gestures have extra communicative value, as they may add meaningful information to what can be derived from the lexico-grammatical structure of an utterance. The problem we want to tackle here is to what extent the nature of these gestures is affected by aspects of the addressee, i.e. (1) the kind of addressee, (2) the addressee's perspective and (3) the relevance of the gesture for the addressee. The various studies presented below were part of a PhD research project of Lisette Mol, in collaboration with her supervisors. More elaborate descriptions of the studies can be found in the respective articles cited below.

## **Study 1: the kind of addressee**

### **Background**

The first study (Mol et al. 2009) looks at the extent to which speakers design their gestures to characteristics of the audience. The idea behind this research relates to the question regarding the extent to which gestures are primarily useful for the speaker, as an important component to facilitate the speech production, or whether they should be viewed as signals that serve a communicative function that is relevant for the addressee.

### **Paradigm**

Participants were invited to watch a Sylvester and Tweety cartoon, and were then asked to recount scenes of the cartoon to another person or to a machine (a presumed audiovisual speech recognizer). The data below come from identical lab settings, where speakers were instructed to speak into a camera, and could not see the addressee, so that possible differences in their gesture and language behaviour could only be attributed to the speakers' beliefs about the kind of addressee.

### **Results**

Participants produced more gestures when addressing a human than a machine, and used relatively few large and pointing gestures to an artificial

addressee. A follow-up perception experiment revealed that judges who were presented with video fragments from the two conditions could tell above chance level whether a clip was from an interaction with a human or machine, suggesting that they indeed expect different interaction styles.

## **Study 2: the addressee's perspective**

### **Background**

The second study (Mol et al 2011) explored whether speakers also take into account the addressee's visual perspective. We investigated whether a speaker would gesture differently, depending on whether speaking partners could or could not see each other. If speakers would gesture differently in contexts where the addressee could or could not see them, this could be interpreted as evidence that the context has an impact on gesture behaviour.

### **Paradigm**

The paradigm was basically the same as the one for study 1, except that this time, we introduced settings whereby the addressee could see the speaker through a webcam (but not the other way around), the speaker could see the addressee (but not the other way around), in addition to conditions with mutual visibility and no visibility (sound only).

### **Results**

Participants produced relatively more and larger gestures when they knew the addressee could see them, even when speakers themselves could not see their addressee, providing evidence that speakers took into account an addressee's perspective when producing gestures. However, note that the webcam condition also led to an unnatural situation as it did not allow direct eye-contact between participants. In a follow-up study, where we used a so-called Eyecatcher which allows direct eye-contact (in that sense being similar to a face-to-face condition), we found that speakers did produce more gestures when seeing the addressee.

## **Study 3: the relevance of the gesture**

### **Background**

The third study (Mol et al 2012) tried to gain more insight into the reasons why people repeat gestures from their partners. The question we wanted to address is whether cases of gestural mimicry should be viewed as manifestations of simple copying behaviour, or whether the repeated gestures are tied to meaning in the linguistic context.

### **Paradigm**

To test this, we conducted a number of experiments, whereby participants would first watch video recordings who would tell scenes of a story to them, which the participant would subsequently have to retell to the experimenter.

### **Results**

A first experiment had two conditions, one where the confederate would sometimes produce gestures vs a condition without any gesturing. Participants more often produced gestures when they had seen them from the confederate. In a second experiment, we compared a condition where the confederate would sometimes produce gestures that did not match the semantic content (e.g. a running gesture when talking about binoculars) vs a condition where the gestures always matched the content. Participants more often repeated gestures if they were meaningful.

### **Conclusion**

This paper presented four studies that together provided evidence that speakers adapt their gestures to the communicative context, in particular characteristics of the conversation partner: their gestures appear to be dependent on the type of addressee, the perspective of the addressee and whether the gesture made sense for an addressee. These addressee-dependent aspects of gesture production suggest that movements of the finger, hand and arms are not solely produced to facilitate a speaker's speech production, but also have a clear functional role to meet the needs of addressees.

### **References**

- Branigan, H.P., Pickering, M.J., Pearson, J., McLean, J.F. 2010. Linguistic alignment between humans and computers. *J. of Pragmatics* 42, 2355–2368.
- House, D., Bell, L., Gustafson, K. Johansson, L. 1999. Child-directed speech synthesis. *Proc. Eurospeech*, Budapest, Hungary, Sept. 1999.
- Kendon, A. 2004. *Gesture: Visible action as utterance*. Cambridge University Press.
- McNeill, D. 2005. *Gesture and Thought*. Chicago: University of Chicago Press.
- Mol, L., Krahmer, E., Maes, A.A., Swerts, M. 2009. The communicative import of gestures. *Gesture* 9, 97-126.
- Mol, L., Krahmer, E., Maes, A.A., Swerts, M. 2011. Seeing and being seen. *Journal of computer-mediated communication* 17, 77-100.
- Mol, L., Krahmer, E., Maes, A.A., Swerts, M. 2012. Adaptation in gesture. *Journal of Memory and Language* 66, 249-264.
- Swerts, M, Krahmer, E. 2010. Visual prosody of newsreaders. *Journal of Phonetics* 38, 197-206.

# L2/L3 acquisition of the English indefinite article

<sup>1</sup>Asma Al-hawi, Sviatlana Karpava<sup>2</sup>

<sup>1</sup>Division of Education & Linguistics, University of Central Lancashire, Cyprus/UK

<sup>2</sup>Department of English studies, University of Cyprus, Cyprus

DOI:

## Abstract

This study aimed to explore the L2/L3 acquisition of the English indefinite article by L1 speakers of Palestinian/Jordanian-Arabic (PJ/A) and Cypriot-Greek (CG) in Jordan and Cyprus. Four groups were recruited: an English control group, two L2 groups, whose L1s are PJ/A or CG, and an L3 group, whose L1 is PJ/A with L2 CG. A forced-choice-elicitation task was used for data collection. Data analysis revealed that the L2 PJ and L2 CG participants had both negative and positive/negative transfer from their L1 into L2, respectively. The results of the L3 participants revealed that their performance was similar to the L2 CG group. Overall, the results of the L2/L3 groups lend partial support to the Fluctuation Hypothesis (FH).

Key words: L2/L3 acquisition, indefinite article, transfer

## Cross-linguistic differences among English, PJ/A and CG

In PJ/A, there is no morphological realization of the indefinite article (Abudaljuh 2016). Therefore, indefinite NPs are always bare. In English and CG, the indefinite article exists. Yet, in Greek singular bare NPs are sanctioned when they are the objects of (i) light verbs or (ii) verbs of accomplishment (Marinis 2003), see examples (1) and (2):

- (1) **ehi** Ø kali dulia (light verb)  
**Has** INDEF.SG– [good job]  
'He has a good job.' (Alexiadou 2014:26).
- (2) **Aghorase** Ø kinito. (verb of accomplishment)  
**buy**–3SG INDEF.SG–cellphone–  
ACC  
'S/he bought a cellphone.' (Kyriakaki  
2011:20).

## Hypotheses and predictions

The data of the study was analysed in light of the predictions of the FH by Ionin et al. (2008). The first prediction is associated with article languages such as CG and; therefore, assumes that transfer overrides fluctuation. Thus, it is expected that the L2 CG participants would not find it difficult to supply

*a(n)* in the target contexts even when they are at lower English proficiency levels. The second prediction is related to article-less languages like PJ/A. In that case fluctuation overrides transfer (ibid). Accordingly, the L2 PJ participants are expected to fluctuate between the definiteness setting and the specificity setting of the article choice parameter until they had enough input to help them fix the parameters associated with the English indefinite article.

Thus, the primary goal of this study is to examine the L2/L3 acquisition of the English indefinite article by seeking answers to the following research questions (RQs):

- R Q1: Is there any L2/L3 English proficiency effect on the performance of each of the L2/L3 groups?  
R Q2: Can the patterns of acquisition of the experimental groups be explained in light of the predictions of the FH (Ionin et al., 2008)?

### **Participants and data collection methods**

Four groups of participants were recruited in Cyprus and Jordan: English native speakers, control group (n=27); two L2 groups: L2 PJ (n=91) and L2 CG (n=93) and an L3 PJ-CG-E group (n=52). Data was collected by means of a written FCET and a questionnaire. The questionnaire aimed to collect some information about the participants. One piece of information that this study is interested in is related to the L2/L3 English proficiency levels of the participants. Accordingly, the participants of the experimental groups were classified into five English proficiency levels. The English levels were as follows: low intermediate (LI), intermediate (Inter), upper intermediate (UI), advanced (Adv) and upper advanced (UA).

Following Ionin, Ko and Wexler (2004), two contexts were examined in the FCET and each context included six sentences. In each sentence, the participants were provided with three options: (the, a/n, Ø). The first context examined NPs with the features [–definite, +specific], see example (3):

- (3) I attended *a workshop* in statistics. It was boring.  
The second context examined NPs with the features [–definite, –specific], see example (4):  
(4) My aunt bought *a house*, but I don't know where exactly.

## Results and discussion

The results showed that English proficiency had an effect on the performance of the L2/L3 groups. In comparison with the EN group, none of the experimental groups was native-like in supplying *a(n)*, see Table 1.

Table 1. Elicitation test results per English proficiency level.

Groups	English proficiency	[+specific]			[-specific]		
		a(n)	the	zero	a(n)	the	zero
L2 PJ	LI	35.2%	24.1%	40.7%	57.4%	22.2%	20.4%
	Inter	59.4%	19.8%	20.8%	71.9%	12.5%	15.6%
	UI	65.3%	18.0%	16.7%	88.0%	3.3%	8.7%
	Adv	74.7%	17.3%	8.0%	86.7%	8.6%	4.7%
	UA	83.3%	8.4%	8.3%	92.7%	5.2%	2.1%
	Total	67.0%	17.1%	15.9%	82.6%	8.6%	8.8%
L2 CG	LI	47.9%	25.0%	27.1%	52.0%	24.0%	24.0%
	Inter	72.6%	14.3%	13.1%	88.1%	4.8%	7.1%
	UI	82.6%	9.1%	8.3%	89.4%	4.5%	6.1%
	Adv	93.9%	2.3%	3.8%	95.5%	0.0%	4.5%
	UA	93.9%	3.5%	2.6%	93.9%	0.0%	6.1%
	Total	80.1%	9.9%	10.0%	85.1%	5.9%	9.0%
L3 PJ-CG-E	LI	46.2%	28.2%	25.6%	56.5%	25.6%	17.9%
	Inter	57.4%	27.8%	14.8%	63.0%	25.9%	11.1%
	UI	75.0%	15.3%	9.7%	80.6%	12.5%	6.9%
	Adv	90.0%	8.3%	1.7%	95.0%	3.3%	1.7%
	UA	93.8%	4.1%	2.1%	93.7%	2.1%	4.2%
	Total	70.5%	17.6%	11.9%	76.3%	14.7%	9.0%
EN	Native	97.5%	0.6%	1.9%	98.2%	0.0%	1.8%

A Pearson correlation coefficient was computed to evaluate the relations between the L2/L3 group's proficiency level and the participants' article use in each context. As an answer to RQ1, the results of the L2 PJ group indicate a positive correlation between the two variables in context E ( $r = .571$ ,  $p = .000$ ) and in context F ( $r = .423$ ,  $p = .000$ ). For the L2 CG group, it was also found a positive correlation between the two variables in context E ( $r = .661$ ,  $p = .000$ ) and context F ( $r = .561$ ,  $p = .000$ ). The results of L3 PJ-CG-E group indicate that there was a positive correlation between the two variables in context E ( $r = .706$ ,  $p = .000$ ) and in context F ( $r = .571$ ,  $p = .000$ ).

Regarding the L2 PJ group's target-like performance, the pair wise comparisons show that the percentages of the participants from the LI ( $p = .022$ ), Inter ( $p = .029$ ), UI ( $p = .000$ ) and Adv ( $p = .047$ ) levels in the [-specific] context were significantly higher than their percentages in the

[+specific] context. Consistent with the predictions of the FH, the pairwise t-tests of the substitution errors reveal that the percentages of the L2 PJ participants' in providing *the* in the [+specific] context were significantly higher than the percentages in [-specific] context ( $p=.010$ ). The specificity effect decreased the more advanced the participants were in English. The pair wise comparisons of the L2 CG group and the L3 groups, on the other hand, showed that neither the use of *a(n)* nor *the* in the [+specific] context were significantly different from the [-specific] context ( $p>0.05$ ).

## Conclusion

Overall, it was found that the L2/L3 participants of higher proficiency in English were more target-like than the participants of lower English proficiency levels. Yet, the results reveal that the L2 PJ participants seemed to have more difficulty in supplying *a(n)* in the [+specific] context than [-specific] context as their L1 lacks the indefinite article. In contrast, the L2 English proficiency had an effect on the performance of the L2 CG participants. This means that the L2 CG participants did not seem to benefit from the presence of the indefinite article in their L1 at the lower English proficiency levels. On the other hand, the L3 advanced and upper advanced participants benefited from their L2 Greek in supplying *a(n)* properly. They were even more target-like than the L2 PJ group of the comparable English proficiency levels. The results of the L2 PJ group supported the FH as the participants showed evidence of fluctuation and proficiency effect. Yet, the results of the L2 CG and the L3 group were not completely in line with the FH. More specifically, the results of the L2 CG participants showed evidence of a proficiency effect though there was no specificity effect.

## References

- Abudalbh, M. 2016. The acquisition of English articles by Arabic L2-English learners: A semantic approach. *Arab World English Journal* 7, 104-117.
- Alexiadou, A. 2014. *Multiple Determiners and the Structure of DPs*. Amsterdam, John Benjamins Publishing Company.
- Ionin, T., Ko, H., Wexler, K. 2004. Article semantics in L2 acquisition: The role of specificity. *Language Acquisition* 12, 3-69.
- Ionin, T., Zubizarreta, M., Bautista-Maldonado, S. 2008. Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua* 118, 554-576.
- Kyriakaki, M. 2011. DETs in the functional syntax of Greek nominals. *Journal of Greek Linguistics* 11, 293-29.
- Mackey, A., Gass, S. 2005. *Second Language Research: Methodology and Design*. Mahwah, New Jersey, Lawrence Erlbaum Associates, Inc.
- Marinis, T. 2003. *Acquisition of the DP in Modern Greek*. Amsterdam, Philadelphia, John Benjamins Publishing Company.

# L1 and L2 skills in Specific Reading Disorder

Georgia Andreou, Vassiliki Tsela, Maria Liakou

Department of Special Education, University of Thessaly, Greece/ Hellenic Open University, Greece

DOI:

## Abstract

The aim of the present study was to investigate morphosyntactic skills in primary and secondary school students with specific reading disorder and their equivalents with good reading skills, in Greek as a first language (L1) and in French as a foreign language (L2). The results of the study revealed significant differences between the two groups on all morphosyntactic tasks in both languages. Our results support the predictive property of the level of the morphosyntactic skills attained in the L1 for the acquisition of the same skills in the L2 for both students with specific reading disorder and typically developing.

Key words: morphosyntactic skills, specific reading disorder, Greek as a first language (L1), French as a foreign language (L2).

## Introduction

Research has shown that students with specific reading disorder (SRD) have limited morphosyntactic skills compared to students with good reading skills (GRS) in both their first (L1) and second/foreign (L2) languages (Andreou & Tsela, 2015; Casalis, Colé, & Sopo, 2004). Those research findings are based on the hypothesis that there is a close relation between the students' skills in their first language and those in their second/foreign language (Sparks, 2012). At the same time, researchers orientate themselves towards the study of cross-language skills transfer (Abu-Rabia & Shakkour, 2014) and the role of the characteristics of each language system (Seymour, Aro & Erskine, 2003).

Research has also shown that phonological awareness seems to play a significant role in the development of decoding skills, while phonological deficits seem to increase the risk of failure in reading acquisition (Hulme & Snowling, 2014). At the same time, various studies have highlighted that non phonological skills, such as morphological and syntactic awareness, can influence and predict reading ability (Carlisle, 2003). More specifically, research, in both alphabetic and logographic languages, has shown that students with SRD have limited morphosyntactic awareness skills in comparison to their peers with GRS irrespective of their phonological awareness level (Casalis et al., 2004).

### **Purpose of the study**

The aim of the present study was to investigate morphosyntactic skills in primary and secondary school students with SRD and their equivalents with GRS, in Greek as a first language and in French as a foreign language. Furthermore, a basic objective was to examine whether morphosyntactic skills in L1 could predict morphosyntactic skills in L2 for both groups.

### **Method**

The sample of the study consisted of one hundred and eight (N=108) students with GRS and one hundred and eight (N=108) students with SRD (average age 12,55 (SD 1,35). Both groups were assessed for their morphosyntactic skills in both languages with four tasks from the standardized for the Greek language tool “Learning Disabilities Reading Inventory” (Padeliadu & Antoniou, 2008) and four equivalent tasks from a non-standardized reading tool for the French language made for the purposes of the study.

The morphosyntactic tasks contained two morphological awareness tasks (verb conjugation and production of compound words) and two syntactic tasks (word order of sentences accompanied by pictures and without the help of pictures).

The comparisons between mean values were made based on the Least Significant Difference (LSD) criterion at a significance level of  $P < 0,05$ . The statistical analyses were made by the SPSS v.15.0 software.

### **Results and discussion**

The results of the study revealed significant differences and correlations between the two groups on all morphosyntactic tasks in both languages: morphosyntactic abilities of students with SRD in both languages were weaker than those of students with good reading skills.

Students with SRD had a statistically significant lower performance than students with GRS in the morphosyntactic tasks in both Greek and French. The low performance in morphosyntax can be attributed to the phonological deficit which students with SRD have, as has been previously shown by other studies (Hulme & Snowling, 2014). The phonological characteristics of morphemes are difficult for students with SRD and, consequently, they face difficulties in phonological processing (Casalis et al., 2004).

Table 1. Performance of students with SRD and with GRS in Greek and French in morphosyntactic tasks.

Group	Grade	Greek		French	
		Mean	SD	Mean	SD
GRS	5 <sup>th</sup> Elementary	93,60	7,95	76,85	9,47
	6 <sup>th</sup> Elementary	94,44	5,41	79,11	7,48
	1 <sup>st</sup> Secondary	95,17	5,05	82,45	6,99
	2 <sup>nd</sup> Secondary	95,22	5,10	84,44	7,84
	3 <sup>rd</sup> Secondary	95,32	5,07	86,57	8,54
SRD	5 <sup>th</sup> Elementary	47,08	7,72	17,90	4,97
	6 <sup>th</sup> Elementary	47,26	8,48	23,11	7,19
	1 <sup>st</sup> Secondary	47,58	6,37	24,80	7,01
	2 <sup>nd</sup> Secondary	47,64	6,62	25,37	7,81
	3 <sup>rd</sup> Secondary	47,92	6,41	29,17	6,18
	MSD <sub>0,05</sub>	4,16		4,16	
GRS Total		94,76	5,81	81,41	8,65
SRD Total		47,48	7,03	23,63	7,49
	MSD <sub>0,05</sub>	1,86		1,86	

GRS: Good Reading Skills, SRD: Specific Reading Disorder, Mean: Mean Value,  
SD: Standard Deviation, MSD: Minimum Significant Difference ( $P<0,05$ )

Table 2. Correlation of morphology and syntax in Greek with the respective skills in French in GRS and SRD students.

Group	Tests in Greek/French	Morphology	Syntax
GRS	Morphology	$r=0,708$	$r=0,699$
		$P<0,001^*$	$P<0,001^*$
	Syntax	$r=0,624$	$r=0,446$
		$P<0,001^*$	$P<0,001^*$
SRD	Morphology	$r=0,248$	$r=0,254$
		$P=0,010^*$	$P=0,008^*$
	Syntax	$r=0,499$	$r=0,393$
		$P<0,001^*$	$P<0,001^*$

GRS: Good Reading Skills, SRD: Specific Reading Disorder,  
\*Statistically Significant difference

In addition, the results of the present study come to agreement with previous research findings which showed that students with SRD, irrespective of their level of phonological awareness, have limited morphological awareness compared to their peers with GRS (Casalis et al., 2004). The morphological tasks in this study also contained words with phonological transparency which may not have been the reason for the difficulties the students faced and their low overall performance. Furthermore, the transparency degree of the Greek language (shallow system) and the French language (deep system) influenced and determined the performances of both groups (Seymour et al., 2003). The unpronounceable endings which characterize the morphology of the French language and the great number of correlation between phonemes and graphemes in French can be especially difficult for students (Ziegler & Goswami, 2005). Finally, our results revealed that the deficits appeared in a morphosyntactic task in one language were transferred analogically to the same task in the other language in both groups, which shows that difficulties in the morphology and syntax in L1 can have an immediate and severe impact on the L2 especially for students with SRD (Abu-Rabia & Shakkour, 2014; Sparks, 2012).

## References

- Abu-Rabia, S., & Shakkour, W. 2014. Cognitive Retroactive Transfer (CRT) of Language Skills among Trilingual Arabic Hebrew and English Learners. *Open Journal of Modern Linguistics*, 4, 1-20.
- Andreou, G., & Tsela, V. 2015. Foreign language by dyslexics in Applied Linguistics: The case of Greek (L1) dyslexic children learning French (L2) as a foreign language. *Studies in Greek Linguistics*, 35, 80-91.
- Carlisle J. F. 2003. Morphology matters in learning to read: a commentary. *Reading Psychology*, 24(3-4), 291-322.
- Casalis, S., Colé, P., & Sopo, D. 2004. Morphological awareness in developmental dyslexia. *Annals of Dyslexia*, 54(1), 114-138.
- Hulme, C., & Snowling, MJ. 2014. The interface between spoken and written language: developmental disorders. *Phil. Trans. R. Soc.* 369, 20120395.
- Padeliadu, S., & Antoniou, F. 2008. *Learning Disabilities Reading Inventory*. Athens: YPEPTH: EPEAEK.
- Seymour, P. H. K., Aro, M., & Erskine, J. M. 2003. Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143-174.
- Sparks, R. (2012). Individual differences in L2 learning and long-term L1-L2 relationships [Supplement]. *Language Learning*, 62, 5-27.
- Ziegler, J. C., & Goswami, U. 2005. Reading Acquisition, Developmental Dyslexia, and Skilled Reading across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, 131, 3-29.

# Discourse in Mild Cognitive Impairment

Malin Antonsson<sup>1</sup>, Kristina Lundholm Fors<sup>1</sup>, Dimitrios Kokkinakis<sup>1,2</sup>

<sup>1</sup>The Swedish Language Bank, Dept. of Swedish, Uni. of Gothenburg, Sweden

<sup>2</sup>Centre for Ageing and Health – AgeCap, University of Gothenburg, Sweden

DOI:

## Abstract

This paper reports on how persons with mild cognitive impairment (MCI) perform on two types of narrative tasks compared to a group of healthy controls (HC). The first task is a widely used picture description task and the other task is a more complex discourse task. Since the latter task puts higher demands on cognitivelinguistic skills, as seen in previous research, we expected this task to be more efficient in discriminating between the two groups. The results confirm this hypothesis.

Key words: mild cognitive impairment, dementia, narrative language

## Introduction

Dementia is a neurodegenerative disorder that is increasing in prevalence as the global population ages. Mild cognitive impairment (MCI) is a condition that is regarded as a stage preceding dementia (Reisberg & Gauthier 2008). Linguistic analysis has shown great potential in identifying early signs of dementia, since the disorder negatively affects language production at all levels (Lancashire & Hirst 2009). Although linguistic features are more successful in identifying persons with clinically manifest dementia than persons with MCI (see e.g. López-de Ipiña et al. 2013), a deterioration in both vocabulary and syntax can be seen several years before a dementia diagnosis (Le et al. 2011). However, differences between healthy controls and persons with MCI are subtle and sometimes difficult to capture (Lundholm Fors et al. 2018; Fraser et al. 2018).

Research has shown that persons with MCI need complex tasks in order for their linguistic difficulties to become evident, as they are able to compensate for their linguistic deficits to some extent (Mueller et al. 2018). Fleming (2014) used a so-called ‘planning task’ to elicit complex discourse. Although this task is suggested to be more sensitive to language decline than a picture elicited narrative, no study has yet to our knowledge compared the two types of tasks for this population. Hence, the purpose of this paper is to present and compare some preliminary results of the linguistic analyses on the two tasks.

## Method

### Participants

A total of 23 patients with MCI and a group of healthy controls (HC;  $n=34$ ) participated in the study. A significant difference between the groups was seen in the Mini Mental State Examination (a test of cognitive functioning), but not on age (mean age in years: MCI: 71.7, HC: 69.2) or education (mean education in years: MCI 13.9, HC 13.3). Detailed inclusion and exclusion criteria are described in Kokkinakis et al. (2017). The study was approved by the local ethical committee review board (reference number: 206–16, 2016).

### Data and statistical analyses

The tasks consisted of a picture description task and a narrative discourse task. The aim of the picture description task was to describe the Cookie theft picture (from BDAE; Goodglass & Kaplan, 1983). The discourse task is modelled on the ‘Trip to New York’ task described in Fleming (2014). In brief, the participants were asked to describe how they would prepare for and execute a trip to Stockholm. The audio recordings were manually transcribed and automatically annotated using Sparv (Borin et al. 2016). Alignment of the audio recordings and transcriptions was made using Webmaus (Kisler et al. 2017), with post-corrections done manually.

Since a recent review (Filiou 2019) suggested outcome measures should include aspects of speech production, fluency and semantics, both tasks were analysed and compared on production length (number of words and number of sentences), pronoun use (pronoun index i.e. pronoun / pronoun + nouns), and some aspects of fluency (ratio of false start to number of words and ratio of fillers to number of words). To investigate differences between participants with MCI and HC we compared the groups using a MANOVA with the above-mentioned variables as dependent variables.

## Results

In the cookie theft task, the groups showed no significant differences ( $F(5, 50)=0.47$ ,  $p=0.80$ ), whereas the groups differed on their performance on the planning task ( $F(5, 50)=2.68$ ,  $p=0.03$ ); see details in table 1. The groups predominantly differed on aspects of productivity, where the MCI group produced fewer words and a smaller number of sentences than the HC group. Although the MCI group had a higher pronoun index and a lower false start ratio than the HC group in the planning task, these differences were not statistically significant. The number of fillers relative to the number of words did not differ between the groups.

Table 1. Comparison of the two groups on Cookie theft and Planning task.

	Cookie theft			Planning task		
	MCI	HC	sig.	MCI	HC	sig.
Words (n)	142.0	172.7	n.s	325.4	426.4	t = 5.8, p = 0.019
Sentences (n)	11.5	12.7	n.s	22.8	30.7	t = 8.18, p = 0.006
Pronoun index	0.36	0.36	n.s	0.55	0.58	n.s
False start ratio	0.005	0.005	n.s	0.006	0.01	n.s.
Fillers ratio	0.02	0.02	n.s	0.03	0.03	n.s.

## Discussion

Whereas the chosen linguistic features did not reveal any differences when comparing the performance of persons with MCI and HC on the cookie theft task, the analysis of the planning task reveals a significant difference between the two groups. This supports the hypothesis that the latter puts higher demands on cognitive-linguistic skills and therefore is better in discriminating between persons with MCI and healthy controls. The persons with MCI produced fewer words and fewer sentences than the HC group, a result similar to Fleming & Harris (2008), that related the shorter length in the MCI group to decreased competence in planning, problem-solving, and organization. Somewhat puzzling this difference was not seen in their more recent study (Fleming 2014) where no difference in length was observed, but in semantic content. The measure used for semantic content in this study, the pronoun index, did not indicate any differences between the groups.

In future studies we will evaluate a larger number of features describing semantic performance, to explore whether differences can be captured with a fine-grained analysis. Furthermore, since the MCI group produced shorter narratives, we plan to analyse the semantic content through thematic coding (i.e. identification of core elements based on Fleming et al. 2008) and frame analysis using Swedish FrameNet++ (Borin et al. 2010).

## Acknowledgements

This work has received support from The Swedish Foundation for Humanities and Social Sciences (grant agreement no: NHS 14-1761:1).

## References

- Borin, L., Dannélls, D., Forsberg, M., Toporowska Gronostaj, M., Kokkinakis, D. 2010. Swedish FrameNet++. Proceedings of 3rd SLTC- Linköping, Sweden.
- Borin, L., Forsberg, M., Hammarstedt, M., Rosen, D., Schumacher, A., Schäfer, R. 2016. Sparv: Språkbanken's corpus annotation pipeline infrastructure. Proceedings of 6th SLTC. Umeå, Sweden.

- Filiou, R-P., Bier, N., Slegers, A. Houzé, B., Belchior, P., Brambati, S. M. 2019. Connected speech assessment in the early detection of Alzheimer's disease and mild cognitive impairment: a scoping review. *Aphasiology*.1-33.
- Fleming, V.B., Harris, J.L. 2008. Complex discourse production in mild cognitive impairment: Detecting subtle changes. *Aphasiology*. 22:7-8, 729-740.
- Fleming, V.B. 2014. Early detection of cognitive-linguistic change associated with mild cognitive impairment. *Comm. Disorders Quarterly*. 35(3), 146-157.
- Fraser, K.C., Lundholm Fors, K., Eckerström, M., Themistocleous, C., Kokkinakis, D. 2018. Improving the sensitivity and specificity of MCI screening with linguistic information. In *RaPID-2@LREC*. Japan.
- Goodglass, H., Kaplan, E., Barresi, B. 1983. *Boston Diagnostic Aphasia Examination*. Philadelphia: Lippincott Williams & Wilkins. W. Kluwer.
- Kisler, T., Reichel U.D., Schiel, F. 2017. Multilingual processing of speech via web services. *Computer Speech & Language*. Vol. 45: 326–347.
- Kokkinakis D., Lundholm Fors, K., Björkner, E., Nordlund, A. 2017. Data collection from persons with mild forms of cognitive impairment and healthy controls - infrastructure for classification and prediction of dementia. *NoDaLiDa*. 131. LiU Electronic Press;172–82.
- Lancashire, I., Hirst, G. 2009. Vocabulary Changes in Agatha Christie's *Mysteries* as an Indication of Dementia: A Case Study. The 19th Rotman Research Institute Conf., *Cognitive Aging: Research and Practice*. Toronto.
- Le, X., Lancashire, I., Hirst, G., Jokel, R. 2011. Longitudinal Detection of Dementia through Lexical and Syntactic Changes in Writing: A Case Study of Three British Novelists. *J of Lit and Ling Computing*. 26(4): 435-461.
- López-de Ipiña, K., Alonso, J-B., Travieso, C.M., Solé-Casals, J., Egiraun, H., Faundez-Zanuy, M., Ezeiza, A., Barroso, N., Ecay-Torres, M., Martinez-Lage, P., Martinez de Lizardui, U. 2013. On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors*. 13(5):6730–6745.
- Lundholm Fors, K., Fraser K.C., Kokkinakis D. 2018. Automated syntactic analysis of language abilities in persons with mild and subjective cognitive impairment. *Studies in health technology and informatics*. 247:705–709.
- Mueller, K.D., Kosciak, R.L., Hermann, B.P., Johnson, S.C., Turkstra, L.S. 2018. Declines in connected language are associated with very early mild cognitive impairment. *Frontiers in ageing neuroscience*. 9:437.
- Reisberg, B., Gauthier, S. 2008. Current evidence for subjective cognitive impairment (SCI) as the pre-mild cognitive impairment (MCI) stage of subsequently manifest Alzheimer's disease. *Int Psychogeriatr*. 20:1–16.
- Szatloczki, G., Hoffmann, I., Vincze, V., Kalman, J., Pakaski, M. 2015. Speaking in Alzheimer's Disease, is that an early sign? Importance of changes in language abilities in Alzheimer's disease. *Frontiers in Aging Neuroscience*. 7, 195.
- Wendelstein, B., Stegmeier, J., Frankenberg, C., Felder, E., Schröder, J. 2015. Changes in the use of pronouns in spoken language in the course of preclinical to early stages of Alzheimer's disease. *Alzheimer's & Dementia: J of the Alzheimer's Association*. 11(7):784-P5.

# Phonosemantic shape-sound correlates

Elena Besedina

Dept of Foreign Languages, Saint Petersburg Electrotechnical University, Russia

DOI:

## Abstract

This paper focuses on the study of sound symbolism and presents the results of two complementary psycholinguistic experiments carried out with 360 male and female Russian native speakers. The data were analysed by statistical methods of experimental data processing, with a relatively large final sample size ( $n \geq 30$ ) and prove the existence of an unambiguous phonosemantic correlation between the rounded shape of an object and the labial phonotype. The obtained results are strikingly consistent with those of a previously conducted structural analysis of the corpus of 95 Russian words denoting ‘roundness’.

Key words: labial phonotype, psycholinguistics, shape-sound correlates, sound symbolism

## Introduction: psycholinguistic experimental data

The association between sound and meaning known as sound symbolism is important for cognitive linguistic theories. Though it concerns the so-called peripheral system of any natural language, it is closely connected with such fundamental problems as arbitrariness/nonarbitrariness of the language sign, presenting a very strong counterargument against the thesis about arbitrary relations between a *signifiant* and a *signifié*, and is highly relevant to the studies of language origin and development.

Throughout the history of studying the issue many researchers who looked at the vocabularies of different natural languages pointed out the special role that certain labial phonemes have in words denoting rounded, spherical, protruding or bulging objects (e.g. Gonda 1940; Ultan 1978; Voronin 2006; Makarov 2017).

Psycholinguistic experimental studies have also proved to be an effective instrument to demonstrate a consistent correspondence between sound and meaning. A certain correlation between labial phonemes and the notion of roundness was also proven experimentally, the classical example being the world-famous ‘maluma-takete’ experiment by W. Kohler (Kohler 1929) which was followed by a number experiments employing similar techniques (Fox Elena Besedina 1935; Ohwaki, Sato 1954; O’Boyle, Tarte 1982; Managhan, Mattock, Walker 2012; Ozturk, Krehm, Vouloumanos 2013).

## Material, method and data procedure

This study aims at proving the existence of a phonosemantic correlation between the rounded shape of an object and the labial phonotype by means of two complimentary psycholinguistic experiments.

In the first experiment the participants (300 male and female native Russian speakers aged 14-24 years) were asked to construct a word for some rounded object according to the CVC model, each giving one and only one response.

The second experiment was carried out with 60 male and female Russian native speakers aged 5-9 years. Due to their specific stage of language development, this age group seemed perfect for that kind of experiment. This time the respondents were asked to “invent” a word for a rounded object made of wax with no restrictions whatsoever. Like in the previous experiment, each child responded only once.

To assess the results statistically, the data obtained in each experiment were analysed by statistical methods of experimental data processing that offer reliable estimates of distribution of a discrete random variable, with a relatively large final sample size ( $n \geq 30$ ). These important parameters include:  $\bar{x}$  – sample mean value of a random variable and its confidence interval,  $S_x$  – (sample) standard deviation,  $S_x / \sqrt{n}$  – Root Mean Square Error (RMSE), where  $n$  is the sample size, and  $v_\mu$  the coefficient of variation of the mean value of a random variable.

## Results and discussion

### First experiment

In the first experiment, the respondents produced 300 CVC words to name a rounded object. The analysis showed that 91% of these words contained at least one labial phoneme, 29 % contained more than one labial phoneme, a quarter of all polylabial words consisting of labial phonemes only. It also interesting that 72 % of all constructed words begin with a labial consonant. The most obvious examples of the words constructed by the respondents are following: боп /'bop/, бот /'bot/, буп /'bup/, бун /'bun/, воб /'vop]/, воф /'vof/, вом /'vom/, мом /'mom/, пом /'pom/ пон /'pon/, фом /'fom/, фоф /'fof/, etc. These words were compared with 412 CVC Russian words obtained by continuous sampling from The Russian Language Dictionary. Further statistical analysis showed that the frequency of the occurrence of the labial phonotype in the analysed words on average exceeds the expectation of this phonotype in Russian words of the CVC model by 1.46 (1.48 for vowels and 1, 43 for consonants). As far as the figures for concrete phonemes are concerned, we can discuss only some of them as the values as the rest of them are not reliable enough to make any suggestions. These

phonemes are /b/, /o/ and /m/ and they exceed the frequency of their occurrence by 2.41; 2.05 and 1.96, respectively.

### Second experiment

In the second experiment, participants produced 60 words to name the rounded object they were shown, and all the words produced contained at least one labial phoneme and 80% of the words were polylabial. The following examples illustrate the results of the second experiment: бофа /'bofa/, вома /'voma/, вуба /'vuba/, курука /ku'ruka/, лума /'luma/, муфи /'mufi/, фом /'fom/, фом /'fom/), чумбак /tʃum'bak/, чучук /tʃu'tʃuk/. Using the same formalism, the comparison of the constructed words and sample of 700 Russian nouns obtained by means of the random number generator revealed that the frequency of the labial phonotype occurrence in the words constructed by young children on average exceeds the expectation of this phonotype in Russian words by 1.73 (2.0 for vowels and 1.54 for consonants). It was also interesting to analyse the relative occurrence frequency of words with labial phonemes. The analysis demonstrated that the frequency of the occurrence of the labial phonotype in these words on average exceeds the expectation of this phonotype in Russian words by 1.43, while for polylabial words this figure is even higher and equals 2.33. The results of these two experiments seem strikingly consistent and show with a 95% confidence level that Russian native speakers exhibit a positive non-arbitrary correlation between the labial phonotype and the rounded shape of an object.

The obtained results correspond well with those of the phonosemantic analysis of the corpus of 95 words denoting 'roundness' the Russian language. When analysed by the same standard statistical methods, they manifested the frequency of labial prototype occurrence that exceeds by 1.91 (2.43 for vowels and 1.61 for consonants) the corresponding variable for the random sample (Besedina 2016).

### Conclusion

The employment of statistical methods not only made apparent the correlation between the designation of 'roundness' and the labial phonotype with Russian native speakers but also demonstrated the efficiency and reliability of this instrument within the framework of this particular problem.

To draw more general conclusions about this association, it is necessary to conduct experiments with groups of respondents of various unrelated languages to find out whether this association between 'roundness' and the labial phonotype is a universalia or at least a frequentialia of sound symbolic character.

## Acknowledgements

The author would like to express sincere appreciation of the assistance in conducting the second experiment provided by the administration of secondary school No. 74 (St. Petersburg).

## References

- Besedina, E. 2016. Sound Symbolism in Designation of Roundness. In: Brodovich, O. (ed.), *Anglistics of the XXI Century. Phonosemantics 2*, 20-24. SPb., Russia.
- Fox, Ch. 1935. An experimental study of naming. *American J. of Psychol.* 47, 545-579.
- Gonda J. 1940. Some remarks on onomatopoeia, sound symbolism and word formation a propos of the theories of C. N. Maxwell. *Tijdschrift voor indischer taal-, land- en volkenkunder* 80, 133-210.
- Köhler, W. 1929. *Gestalt psychology*. NY: H. Liveright, 1929.
- Makarov, A. A. 2017. Kompjutornyi analiz raspredeleniya fonemotipov v slovakh oboznachayustchikh 'okrugloye' i 'bolshoye' (Компьютерный анализ распределения фонемотипов в словах, обозначающих «округлое» и «большое»). In: Flaksman, M. (ed.), *Proceedings of the VI Interacademic Conference 'Current Issues of Linguistics'*, 296-302. SPb., Russia.
- Managhan, P., Mattock, K., Walker, P. 2012. The role of sound symbolism in word learning. *J. Exp. Psychol.: Learn. Mem. Cogn.* 38, 1152-1164.
- O'Boyle, M., Tarte, R. 1980. Implications for phonetic symbolism: the relationships between pure tones and geometric figures. *J. of Psychol. Research* 9, 535-543.
- Ohwaki, Y., Sato, K. 1954. An experimental study on the psychological relationship between visual and auditory stimuli. *Tonoku Psycholinguistic Folia* 28, 58-71.
- Ozturk, O., Krehm, M., Vouloumanos, A. 2013. Sound symbolism in infancy: evidence for sound-shape cross-modal correspondences in 4-month-olds. *J. Exp. Child Psychol.* 114, 173-186.
- Voronin, S. V. 2006. *Osnovy fonosemantiki (Основы фоносемантики)*. М.: ЛЕНАНД.
- Ullian, R. 1978. Size-sound symbolism. *Universals of human language* 2, 523-568.

# Tonology and sentence structure in Greek

Antonis Botinis<sup>1</sup>, Athina Kontostavlaki<sup>1</sup>, Olga Nikolaenkova<sup>2</sup>, Charalambos Themistocleous<sup>3,4</sup>

<sup>1</sup>Lab of Phonetics and Computational Linguistics, University of Athens, Greece

<sup>2</sup>Department of General Linguistics, Saint Petersburg State University, Russia

<sup>3</sup>Department of Neurology, Johns Hopkins University, USA

<sup>4</sup>Department of Swedish, University of Gothenburg, Sweden

DOI:

## Abstract

This is an experimental investigation of tonal production as a function of lexical stress variability and sentence types in Greek. One production experiment was carried out according to which the lexical stress assignment in the last words varied in the last three syllables in the context of polylectic and monolectic statement and alternative (polar) question sentence contexts. The results indicate that statements are associated with a rise-fall tonal pattern encompassing the last stressed syllable whereas the questions are associated with a rise-fall tonal pattern encompassing the last syllable of the utterance. The initial rise of the rise-fall in statements may be truncated whereas the rise-fall in alternative questions remained fairly invariable.

Keywords: tonal production, sentence structure, prosody, phonetics, Greek

## Introduction

This is part of a larger experimental investigation with regards to tonological structures as a function of lexical stress focus and sentence type variability. In the present paper, we will report the results of a qualitative analysis with reference to tonal patterns in variable lexical stress and sentence type contexts. The lexical stress is constrained by a three-syllable window in Greek, namely only ultimate, penultimate and antepenultimate syllables can be associated with stress. At the postlexical level, stress is manifested differently depending on the sentence types and variable focus and sentence types structures. Focus productions in declarative contexts are usually associated with a rise-fall expansion of the tonal range locally on the stressed syllable of the speech unit in focus and a tonal compression on the post focus speech material (Botinis 1989). Alternative questions, usually referred to as “total”, “polar” and “yes/no” questions, are usually associated with a tonal rise-fall pattern at the right edge of the respective speech unit (Botinis, Chaida, Nikolaenkova, Nirgianaki 2016).

This study aims to show the distinct patterns of stress, number syllables and their effects with the distinct tonal patterns of questions and statements in Greek.

## Methodology

The experimental material was recorded in a soundproof room at the University of Athens. A Zoom H4n audio recorder was used, and the recordings were coded as single mono sounds (sampling frequency: 22050 Hz) and transferred to a PC for further analysis using a compact SD card. Tokens were typed in Standard Greek orthography. Test materials were manually segmented and labelled by using simultaneous inspections of waveforms and wide-band spectrograms (see Figures 1-2).

## Results

The results of the present investigation are shown in figures 1-2. Figure 1 shows the production of the Subject-Verb-Prepositional Phrase sentence. Each production has a different stress pattern associated with the final word, namely the stress is in the antepenultimate (Figure 1, upper left and right panel), the penultimate (Figure 1, left and right middle panel) and ultimate (Figure 1 lower left and right panels). All left panels depict statements and all right panels depict Wh. Question tonal patterns.

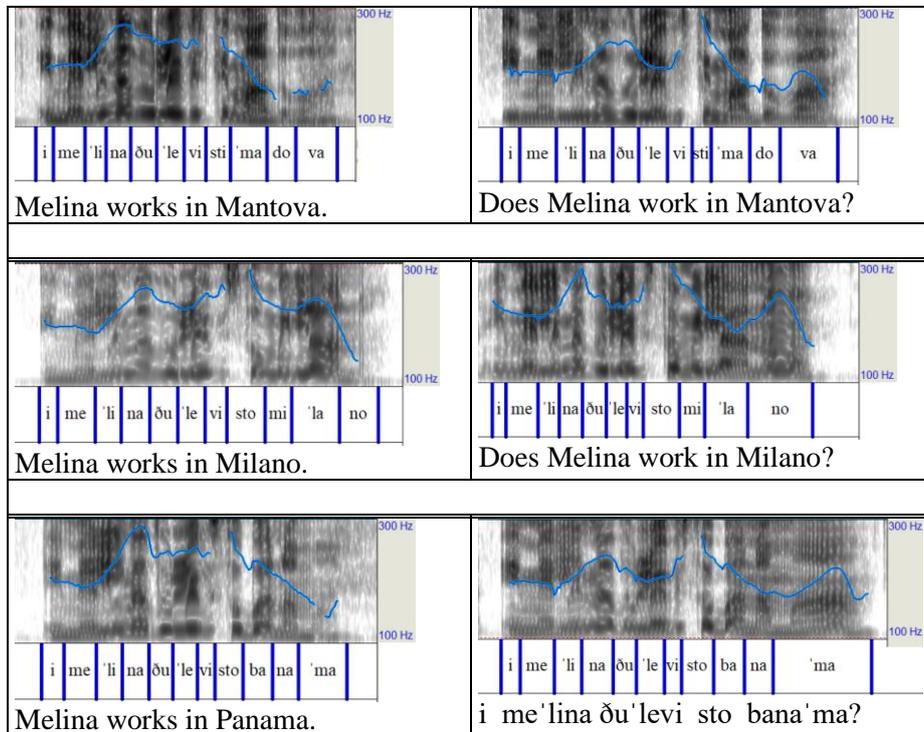


Figure 1. Tonal production of polylectic sentences as a function of lexical stress variability and statement (left) vs. question distinctions (right).

Figure 2 shows the tonal production of statements (Figure 2 left column) and questions in one word examples (Figure 2 right column). Each word production has a different stress pattern associated with the final word, namely the stress is in the antepenultimate (upper panels), the penultimate (middle panels) and ultimate (lower panels).

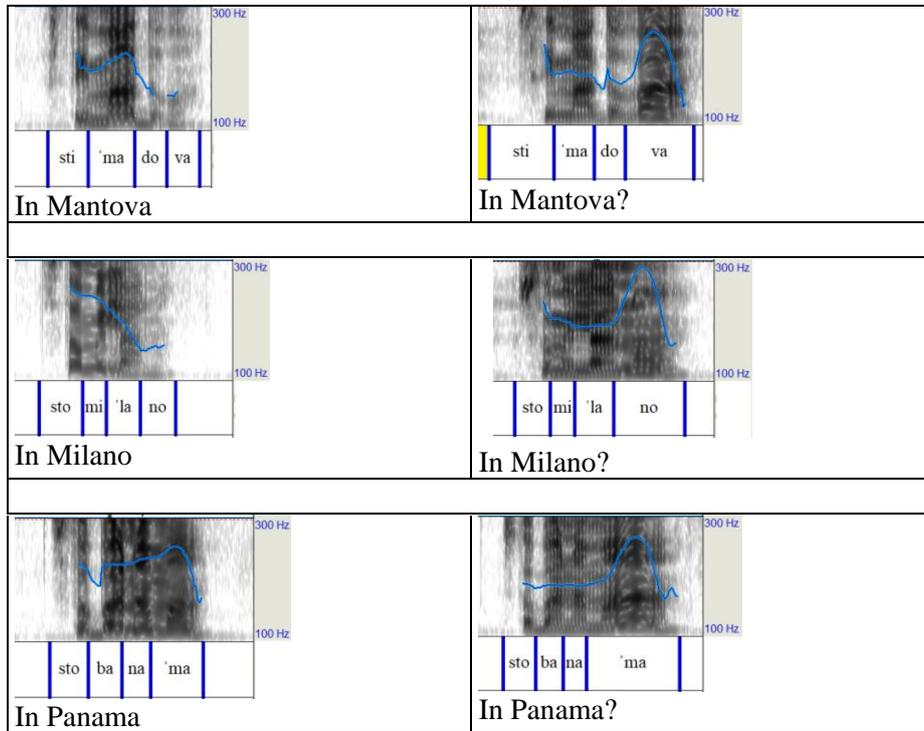


Figure 2. Tonal production of monolectic sentences as a function of lexical stress variability and statement (left) vs. question distinctions (right).

Statements and questions have different tonal patterns. The pattern for statements is a rise on the stressed syllable and a fall at the end of the utterance. This pattern is not evidenced due to constraints imposed by the positioning of the stress. For example, when the stress is on the ultimate syllable this fall does not take place. In questions the pattern is a lowering of the fundamental frequency on the stressed syllable and a rise. The lowering is not evidenced when the stress is on the ultimate syllable, since it coincides with the final rise.

## Discussion

The current study displays that Standard Greek like many other speech varieties exhibits language specific tonal patterns that characteristically distinguish Standard Greek from other language families such as Italian or Swedish and even other language varieties of Greek, such as Cypriot Greek (cf. Themistocleous 2011, 2014). On the other hand, it is related to earlier research in English that displays the distinct tonal patterns between statements and questions especially with the tones that demarcate phrase and sentence boundaries (Beckman, Edwards 1990).

The application of the final movements takes place primarily at the final syllables. In addition, this study demonstrates the effects of edge-tones and prosodic domain boundaries on final lengthening. Specifically, it shows that two types of tonality apply at the right edge of post-lexical domains: (a) pre-boundary tonal movement that marks the edges of prosodic domains and varies depending on boundary strength and (b) edge-tone lengthening that associates with the effects of syllable lengthening. The two types of tonal structure trigger different overlaying effects that account for the lengthening patterns at the edges of prosodic domains.

Beyond the results of the present study, our investigation has demonstrated a straightforward way for the elicitation of alternative questions in a question-question methodological paradigm. This methodology may be applied in a variety of languages with different prosodic and intonation systems, the results of which may be cross-linguistically and cross-dialectically compared in a direct way without any other linguistic variability but prosody.

## References

- Beckman, M., Edwards, J. 1990. Lengthenings and shortenings and the nature of prosodic constituency. In Kingston, J., Beckman, M. (Eds.), *Papers in laboratory phonology I. Between the grammar and the physics of speech*, 152–178. Cambridge University Press.
- Botinis, Antonis. 1989. *Stress and prosodic structure in Greek*. Lund University Press.
- Botinis, A., Chaida, A. Nikolaenkova, O., Nirgianaki, E. 2016. Intonation and polar questions in Greek revisited. In Botinis, A. (Ed.), *Proc. ExLing 2016*, 41-44.
- Themistocleous Ch. 2011. *Prosodia kai plirophoriaki domi stin Ellinici (Prosody and Information Structure in Greek)*. PhD Thesis, University of Athens, Greece.
- Themistocleous Ch. 2014. Edge-Tone Effects and Prosodic Domain Effects on Final Lengthening. *Linguistic Variation* 14(1). 129–160. DOI: 10.1075/lv.14.1.06the.

# Grammatical mood and ambiguity aversion

Sacha Bourgeois-Gironde<sup>1</sup>, Alda Mari<sup>1</sup>, David Nicolas<sup>1</sup>, David Blunier<sup>2</sup>

<sup>1</sup>Institut Jean Nicod, CNRS/ENS/EHESS/PSL, France

<sup>2</sup>Département de Linguistique, Université de Genève, Switzerland

DOI:

## Abstract

This paper explores the impact of language on behaviour studying the impact of mood in a probabilistic choice context. Building on the idea of “aversion to ambiguity” according to which subjects prefer situations in which probabilities are known over those in which they are not known, we have systematically tested the association of sentence mood with choice situations. We found that with indicative the aversion to ambiguity is confirmed, whereas with the subjunctive, it is not. This indicates that grammatical features can influence the way in which subjects apprehend choices in probabilistic contexts.

Key words: decision theory, language, ambiguity, risk, mood, subjunctive

## Goal

While modern economic theories have soundly established a connection between economic behavior and psychology, there is a new and growing interest in the connection between language and behavior.

This paper studies one aspect of this connection by addressing choices in probabilistic contexts. Specifically, we study how the verbalization of mental states accompanying choices in probabilistic contexts can modify what psychology based economic theories have labelled as the “standard” behavior. We show here that certain grammaticalized features of natural language can impact the expectations solely based on psychological considerations.

## Background on decision theory

Decision theory models a relation of preference over a set of options. Typically, it models future oriented decisions, whereby different possible situations are envisaged with different probabilities of being realized. One major distinction established by economist is the one between risk and ambiguity. Consider the ballot box on the right. A situation is pictured, where there are 100 balls, 50 orange and 50 grey. The proportion between the orange and the grey balls is known. The left ballot box pictures a situation of **risk**, where the probabilities are known. The right ballot box represents a situation of **ambiguity**. The ballot box contains 100 orange and grey balls, but the proportion between the two is not known. According to

the view of Savage 1954, subjects should be indifferent to this difference insofar as, subjectively, the two situations are equivalent. This, insofar as the so-called “rational” individual is considered. However, it has been widely shown since Ellsberg 1961 that individuals do not treat these decision contexts as equivalent and display what has been labeled an “aversion to ambiguity” (Epstein 2004). The preference for precise probabilistic information over ambiguous information leads to standard preferences for the former choices.

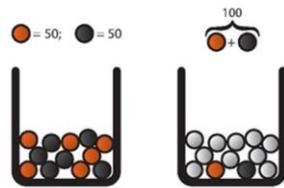


Figure 1. The urn on the left represents a *risky* situation; the urn on the right an *ambiguity* situation.

### The role of language and mood

Some correlations between mood and decision behavior (investment under risk) have been probed at a macro level (Bernhofer, Constantini, Kovacic, 2018). Our study is the first to investigate how linguistic reports about choices intuitively point towards ambiguity averse or ambiguity non-averse in a stylized experimental setting.

In our study we systematically tested the association of sentence mood with choice situations and we have asked the question of whether the “standard” behavior is observed no matter what verbalization of the mental states we use. Specifically, we have studied choices for action in the context of belief predicates.

- (1) I believe that it is raining.

We have used Italian, as this same English sentence can be rendered via two different grammaticalizations of mood on the verb in the embedded clause (called ‘prejacent’)

- (2) Credo che piove.IND/piova.SUBJ -- I believe that it is raining.

In the linguistic literature on mood, there is a widespread consensus for admitting that mood is an indicator of the mental state described by the attitude, in our specific case, the belief state (Portner 2018 for an overview). Indicative is an indicator of uncertainty (this is rendered via a notion of homogeneity in the internal state (Giannakidou, Mari 2010 for a

comprehensive discussion of the literature), subjunctive is an indicator of uncertainty (this is rendered as a non-homogeneity in the mental state). As a result, the use of subjunctive under the belief in \ref{subj} indicates that there is a weakness in the mental belief state of the attitude holder that includes the possibility that both the prejacent and its negation are possible candidates for being true in the actual world. Specifically for Italian 'believe' predicate *credere*, different authors locate at different levels of the interpretation of the semantic component of uncertainty for belief. According to some of them (e.g. Homer 2008) belief is the weaker member of the scale <certainty, belief> and thus non-homogeneity is asserted. According to some others (Mari 2016), the relevant scale is <knowledge, belief>. In this case, uncertainty would be coded in the presupposition with 'believe' bearing the implicated presupposition 'not know' (Mari, 2016, Giannakidou, Mari 2020). No matter whether uncertainty is located in the assertion or in the presupposition, what matters for the purpose of our experiment is that subjunctive triggers *uncertainty*. Our hypothesis is in (3):

- (3) Different verbalizations have an impact on the 'standard' behavior (H<sub>0</sub>)

## Experiment

We ran a questionnaire among native northern- and central-Italian speakers ( $n=64$ , ages between 18 and 50) consisting of 16 randomized items; each participant would see one condition out of four, for a total of 4 questions, corresponding to 4 different choice situations. Each participant was presented a description of a probabilistic choice situation, with two illustrations for each situation: one depicting a **risk** situation, the other depicting an **ambiguity** situation.

Each scenario was accompanied by a target sentence describing the situation. Each sentence would consist of the matrix predicate *credere*, and an embedded verb displaying verbal agreement with either the subjunctive or the indicative mood.

As shown, participants were asked to choose which situation the target sentence illustrated best.

## Results and discussion

We found that with indicative, H<sub>0</sub> is confirmed. With subjunctive, H<sub>0</sub> is not confirmed (p-value: 0.007).

This indicates that, when the indicative is used, the default preference for the "standard" behavior for risky situations is maintained. However, this preference seems to be overridden when the subjunctive is used.

We can explain the results by appealing to the “weakness” of the subjunctive. As we noted above, the subjunctive triggers (a presupposition of) uncertainty. The difference between the risk ballot box and the ambiguity ballot box is that the probabilities are not known for the latter. In other terms, the ambiguity ballot box allows to accommodate uncertainty insofar as there is a parameter (the distribution of the probabilities) unknown to the subject. The subjunctive orientates the choice towards the ambiguity ballot box thus impacting on the standard behavior.

We thus acknowledge that, while the “standard” individual prefers risky situations over ambiguous ones, we see that by making explicit the states of mind that accompany a certain choice and by manipulating grammatical parameters we do no longer obtain the expected generalizations.

This result lead us to open several questions for future research; in particular, we will study whether and to what extent similar results apply to languages that do not have mood shift with belief predicates.

## References

- Bernhofer, J., Constantini, F., Kovacic, M. 2018. Risk attitudes, investment behavior and linguistic variation. University Ca’Foscari, Dept. of Economics Research Paper Series, 34/15.
- Ellsberg, D. 1961. Risk, ambiguity, and the savage axioms. *The quarterly journal of economics*, 643–669.
- Epstein, L.G. 2004. A definition of uncertainty aversion. In *Uncertainty in Economic Theory*. Routledge, 187–224.
- Giannakidou, A., Mari, A. 2020. Veridicality in Grammar and Thought: modality, propositional attitudes and negation. The University of Chicago Press.
- Homer, V. 2008. Intervention effects: the case of presuppositions. Master’s thesis, ENS.
- Mari, A. 2016. Assertability conditions of epistemic (and fictional) attitudes and mood variation. In *Semantics and Linguistic Theory* 26, 61–81.
- Portner, P. 2018. *Mood*. Oxford University Press, 2018.
- Savage, L.G. 1954. *The foundations of statistics*. Courier Corporation.

# Affective analysis of customer service calls

Vera Cabarrão<sup>1,2</sup>, Mariana Julião<sup>1,3</sup>, Rubén Solera-Ureña<sup>1</sup>, Helena Moniz<sup>1,2,4</sup>,  
Fernando Batista<sup>1,5</sup>, Isabel Trancoso<sup>1,3</sup>, Ana Isabel Mata<sup>2</sup>

<sup>1</sup>Laboratório de Sistemas de Língua Falada, INESC-ID Lisboa, Portugal

<sup>2</sup>FLUL/CLUL, Universidade de Lisboa, Portugal

<sup>3</sup>Instituto Superior Técnico, Universidade de Lisboa, Portugal

<sup>4</sup>Unbabel Lda., Portugal

<sup>5</sup>Instituto Universitário de Lisboa (ISCTE-IUL), Portugal

DOI:

## Abstract

This paper presents an affective and acoustic-prosodic analysis of a call-center corpus (700 phone calls with corresponding customer satisfaction levels). Our main goal is to understand how customers' satisfaction correlates to the acoustic-prosodic and affective information (emotions and personality traits) of the interactions. A subset of 30 calls was manually annotated with emotions (frustrated vs. neutral) and personality traits (*Big-Five* model). Results on automatic satisfaction prediction from acoustic-prosodic features show a number of very informative linguistic knowledge-based features, especially pitch and energy ranges. The affective analysis also provides encouraging results, relating low/high satisfaction levels with the presence/absence of customer frustration. Concerning personality, customers tend to express signs of anxiety and nervousness, while agents are generally perceived as extroverted and open.

Key words: acoustic-prosodic features, emotions, personality traits, call-center interactions, customer satisfaction

## Introduction

Customer support is essential to assuring the bonds between customers and service providers, and for maintaining frequent contacts and offering solutions to customers' problems. Measuring customers' satisfaction is necessary to distinguish between satisfied and loyal clients vs. upset ones who will most probably subscribe to services in other companies. Customers' satisfaction correlation with linguistic features (lexical, acousticprosodic) and paralinguistic events (e.g., emotions, personality traits) in customer-agent interactions can help companies to develop automatic procedures to understand whether the customer is satisfied with the service or not, and how the agent performed to achieve a successful outcome. Literature on customer satisfaction (e.g., Lotko, 2015; Delcourt et al., 2013; Stoyanchev, Maiti and Bangalore, 2019), and emotions detection (e.g., Lee and Narayanan, 2005; Vidrascu and Devillers, 2007) has shown diverse findings on how agents interact with customers successfully. Our

main goal is to understand how customers' satisfaction level (satisfied vs. nonsatisfied) correlates with the acoustic-prosodic clues extracted from speech interactions in a European Portuguese (EP) call-center corpus, as well as with emotions and personality traits pertained to either the customer or the agent. To the best of our knowledge, this is the first study on speech-based affective computing for customer satisfaction in EP.

## Methodology

The CCS (call-center satisfaction) corpus comprises around 700 call-center phone calls between customers and agents. The conversations were collected from a non-technical support service in Portugal. All the calls are labelled according to the Net Promoter Score (NPS, Reichheld, 2003), a metric used to express customers' satisfaction. NPS scores were provided by the customers at the end of the calls according to their overall satisfaction with the service and their hypothetical leaning towards recommending it (promoters) or not (detractors) to their friends and family.

This corpus was divided into train, development, and test subsets, and was used for creating a machine learning model for automatic satisfaction prediction. In addition, an expert annotator labelled a subset of 30 calls. Customers' and agents' speech turns were annotated in terms of emotions as either neutral or frustrated, and their overall personality was annotated according to the *Big-Five* traits: openness (artistic, imaginative), conscientiousness (organized, efficient), extroversion (outgoing, talkative), agreeableness (kind, sympathetic), and neuroticism (anxious, worrying).

The experiments here reported use two different sets of acousticprosodic features typically adopted in computational paralinguistic tasks, namely the feature set used in Schuller et al. (2012) and the eGeMAPS set (Eyben et al., 2015), and also a set of knowledge-based features (KB) that take into account linguistic information at the intra and inter-pausal unit levels (IPUs).

## Results

The manual annotation of emotions shows that "neutral" is the most common emotion in both high and low customers' satisfaction groups. This was expected, given the nature of the corpus: in most phone calls, clients aim to solve small problems, and do not tend to present an aggressive posture towards the agent. "Frustration" is mainly associated with customers and usually decreases along the call. The comparison between detractors and promoters does not reveal a significant difference in "frustration" for both groups (22.5% vs. 16.3%, respectively). Therefore, we can hypothesize that, even though frustrated clients are present in both satisfaction groups, the performance of the agent may influence the customers' evaluations towards a high or a low score according to the interactions between them.

For personality traits, the manual annotation of the corpus shows that the most salient traits for the agents in the low satisfaction group are conscientiousness (all the agents), agreeableness (83%), and extroversion (67%). No agent in this group was labelled as neurotic, as opposed to the users (all users were associated with this trait). Clients were also labelled as conscientious (83%). These results show that the most distinctive traits between agents and users in this group are neuroticism and openness (no users were associated with the openness trait). In the group with high satisfaction level, neuroticism and openness are also the most distinctive traits between clients and agents: 5% of the agents were labelled as neurotic vs. 73% of the users, and 83% of agents as open vs. 28% of users. For the remaining traits, they tend to be similar.

Concerning the automatic prediction of the satisfaction levels based on acoustic-prosodic cues, the model (a linear SVM) that uses the linguistic KB-features (41 features only) achieves a satisfactory performance. These features are very informative (unweighted average recall –UAR– of 75.4%), with a better recall for non-satisfied (81%) than for satisfied (68%) users. Combined with the eGeMAPS features, the model improves the recall for satisfied users to 75%. We also carried out a relevance analysis of the KBfeatures based on their corresponding weights in the SVM classifier. Pitch range and dynamics play a major role in the prediction of satisfaction. A possible explanation for these results is related to the linguistic KB-features focusing on dynamics between inter-pausal units (IPUs), which can be related with the customer-agent interactions changing throughout the call. Energy dynamics and range are also informative, helping for distinguishing between satisfied and non-satisfied clients.

## Conclusions

Customer support analytics has evolved due to the impact of affective computing research. The subjective assessment of customer satisfaction is being further supported by objective and quantifiable measures of emotions and other social human behaviors expressed in speech.

Emotions are crucial to better understand the dynamics of customer-agent interactions. Our results show that neutral and frustrated speech segments are transversal in the corpus, but frustrated segments in the promoters group appear usually at the beginning of the call, changing their dynamics influenced by the agent's behavior throughout the conversation. Personality traits do not seem to have a great impact on the satisfaction level discrimination, but are crucial to distinguish customers from agents. To this respect, customers are mainly described as nervous and anxious (neurotic) while agents are perceived as conscious, open and extroverted. Regarding the automatic prediction of the satisfaction levels based on acoustic-prosodic

clues, our models achieved satisfactory performances. Results show that the linguistic KB-features are very informative (UAR of 75.4%), with a better recall for non-satisfied users (81%) than for satisfied. The most informative features are pitch and energy range and dynamics between IPU, which can be associated with customer-agent interactions changing throughout the call. Although these are preliminary results due to the size of the corpus, our results may still be regarded as encouraging.

## Acknowledgements

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with references UID/CEC/50021/2019 and UID/LIN/00214/2019, and under PhD grants SFRH/BD/96492/2013 and SFRH/BD/139473/2018.

## References

- Delcourt, C., Gremler, D.D., Van Riel, A.C. and Van Birgelen, M., 2013. Effects of perceived employee emotional competence on customer satisfaction and loyalty: The mediating role of rapport. *Journal of Service Management*, 24(1), 5-24.
- Eyben, F., Scherer, K.R., Schuller, B.W., Sundberg, J., André, E., Busso, C., Devillers, L.Y., Epps, J., Laukka, P., Narayanan, S.S. and Truong, K.P., 2015. The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. *IEEE Transactions on Affective Computing*, 7(2), 190-202.
- Lee, C.M. and Narayanan, S.S., 2005. Toward detecting emotions in spoken dialogs. *IEEE Transactions on Speech and Audio Processing*, 13(2), 293-303.
- Lotko, A., 2015. Classifying customers according to NPS index: cluster analysis for contact center services. *Central European Review of Economics & Finance*, 8(2), 5-30.
- Reichheld, F.F., 2003. The one number you need to grow. *Harvard Business Review*, 81(12), 46-55.
- Schuller, B.W., Steidl, S., Batliner, A., Nöth, E., Vinciarelli, A. and Burkhardt, F., 2012. The INTERSPEECH 2012 Speaker Trait Challenge. *Proc. of Interspeech 2012*, Portland, OR, USA.
- Stoyanchev, S., Maiti, S. and Bangalore, S., 2019. Predicting interaction quality in customer service dialogs. In *Advanced Social Interaction with Agents*, 149-159. Springer, Cham.
- Vidrascu, L. and Devillers, L., 2007. Five emotion classes detection in real-world call center data: the use of various types of paralinguistic features. *Proc. of International Workshop on Paralinguistic Speech between Models and Data*, Saarbrücken, Germany.

# Interactions of voice quality settings

Zuleica Camargo<sup>1</sup>, Albert Rilliard<sup>2</sup>

<sup>1</sup>Integrated Acoustic Analysis and Cognition Laboratory (LIAAC), Pontifical Catholic University of São Paulo, Brazil

<sup>2</sup>Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI), University Paris-Saclay, France

DOI:

## Abstract

This paper aims at investigating the possible interactions of the voice quality settings (VQS) related to vocal tract (supralaryngeal) and phonatory (laryngeal) activities in the Vocal Profile Analysis Scheme (VPAS), based on Laver's model. The corpus was composed of semi-spontaneous speech samples and repetitions of key-sentences (designed for voice quality evaluation) produced by 71 speakers. The database was perceptually annotated by means of the VPAS, adapted to Brazilian Portuguese (BP), indicating the vocal tract and phonatory VQS. Hierarchical clustering analysis was run. Results reinforced the interactions between the vocal tract and phonatory VQS, having the general muscular tension (vocal tract and laryngeal hyper and hypofunction) adjustments explained the main grouping of those VQS.

Key Words: phonetics, voice quality, auditory perception

## Introduction

Since voice quality has linguistic, paralinguistic and extra-linguistic functions, many approaches are available to the evaluation in terms of its perceptual, acoustic and physiological characteristics.

Classifying voice qualities from the perceptual perspective is still a difficult task and relying on a phonetic description of voice quality as provided by Laver's model (1980) and on the perceptually based Vocal Profile Analysis Scheme, the VPAS (Laver, MackenzieBeck, 2007), have proven useful for many fields, covering crosslinguistic and sociolinguistic variants, clinical demands, forensic and speech technologies applications, and expressive effects in speaking and singing voices (French et al, 2015; Mackenzie-Beck, Schaeffler, 2015; Rilliard, d'Alessandro, Evrard, 2018).

The adaptation of the system was carried out in Brazilian Portuguese context, referred to as the BP-VPAS (Camargo, Madureira, 2008). The basic analytical unit is the voice quality setting (VQS), a long-term muscular tendency in the vocal apparatus (vocal tract and phonatory levels). This system overcomes a traditional limitation of voice quality evaluations: the impressionistic descriptions.

Some of our previous investigations also challenged the demands in processing statistical analysis for the parameters generated by the VPAS annotations (Camargo et al, 2015). For the moment, we are concerned with the interactions between the elements in VPAS system, trying to find instances of correlation among the VQS. This paper addresses the possible interactions of the VQS related to vocal tract (supralaryngeal) and phonatory (laryngeal) levels.

## Methods

The corpus was composed of semi-spontaneous speech samples and repetitions of 10 key-sentences (designed for voice quality evaluation) read by 71 speakers (52 female and 19 male, ranging from 17 to 58 years old). The database was perceptually annotated by means of the BP-VPAS by two expert judges, generating the vocal profiles (the combination of VQS) for 1,608 utterances.

The VPAS annotations of the collected utterances may be considered as a large contingency table, counting the occurrences of each feature. This shows it is a complex set of data, especially because many dimensions of VPAS are, by construction, uncorrelated. However, some relations exist between profiles – especially because the speakers tend to recruit several VQS together, generating compound VQS (Laver, 1980, Laver, Mackenzie-Beck, 2007). This is captured by the analysis presented hereafter.

A correspondence analysis was run on the (raw) VPAS annotations for all speakers and all speaking styles (semi-spontaneous and read). The 20 first dimensions of the analysis explained more than 80% of the variance. They were used as an input for a hierarchical clustering procedure. The number of clusters was determined so as to minimize a criterion of inertia gain that leads to 11 categories built on the patterns of VPAS features, just exposed on results.

## Results and discussion

Results reinforced the importance of interactions between vocal tract and phonatory VQS, specially by the indication of the 11 clusters generated: Cluster #1 (n=15): Raised Tongue Body and Harsh Voice VQS; Cluster #2 (n=22): Advanced Tongue Body, Raised Larynx and Falsetto VQS; Cluster #3 (n=665): Retracted Tongue Body, Retracted Tongue Tip, Laryngeal Hyperfunction and Harsh Voice VQS; Cluster #4 (n=79): Labiodentalization, Minimized Range of Lips and Tongue Body, Pharyngeal Expansion, Vocal Tract Hypofunction and Creaky VQS; Cluster #5 (n=15): Audible Nasal Air Escape and Nasal VQS; Cluster #6 (n=409): Advanced Tongue Tip, Minimized Range of Tongue Body and Jaw, Nasal, Denasal,

Laryngeal Hypofunction, Breathy Voice and Modal VQS; Cluster #7 (n=35): Spread Lips, Extended Range of Tongue Body, Minimized Range of Jaw and Creaky Voice VQS; Cluster #8 (n=123): Spread Lips, Closed Jaw, Lowered Tongue, Pharyngeal Constriction, Vocal Tract Hyperfunction VQS; Cluster #9 (n=144): Advanced and Lowered Tongue Body, Open Jaw, Vocal Tract Hypofunction, Lowered Larynx, Creaky and Breathy Voices VQS; Cluster #10 (n=63): Pharyngeal Expansion, Lowered Larynx and Modal VQS; Cluster #11 (n=36): Lowered Tongue Body, Extended Range of Jaw and Lips, Vocal Tract Hyperfunction and Breathy Voice VQS.

The Clusters #3, #6, #9 and #8 were the largest ones and reinforced the links between vocal tract and phonatory mechanisms, by respective muscular tension settings, in which: laryngeal hyperfunction was combined with vocal tract hyperfunction VQS (pharyngeal constriction, retracted tongue body, raised larynx) and, instead, laryngeal hypofunction was combined with vocal tract hypofunction VQS (advanced tongue tip, open jaw, pharyngeal expansion, lowered tongue body and larynx).

Laver's general muscular tension settings (hyper-hypofunction scale) seemed to explain the main grouping of vocal tract and phonatory VQS. Tense VQS were also related to non-modal (harsh voice) and lax VQS were closely related to modal or creaky voices.

These findings tended to associate an essential element linking vocal tract and phonatory voice events, i.e., the vertical position of the larynx (Fink, Basek, Epanchin, 1956). In hyperfunction muscular VQS, raised larynx was found and on hypofunction VQS, lowered larynx was registered. These indications were drafted in some of our previous explorations of VQS.

Additional remarks were related to tense falsetto linked to raised larynx and advanced tongue body (#2). The findings were related to a recurrent female setting characterized by anteriorization supralaryngeal articulations and breathy voice (#6); sounding "smaller", cf. cliché of "young girl" (Léon 1993). Some sources of variation of VQS are related to the organic structure of the vocal apparatus (the anatomy) and the phonetic usages (Laver, MackenzieBeck, 2007).

## Conclusion

Results reinforced the interactions between vocal tract and phonatory VQS, based on muscular tension adjustments related to hyper and hypofunctional states.

## Acknowledgements

A CNPq grant- PQ (302602/2016-0) is acknowledged.

## References

- Camargo, Z., Madureira, S. 2008. Voice quality analysis from a phonetic perspective: Vocal Profile Analysis Scheme Profile for Brazilian Portuguese (BP-VPAS). In: Barbosa, P.A., Madureira, S.; Reis, C. (Ed.), Proc. 4<sup>th</sup> Conference on Speech Prosody, 57-60, Campinas, Brazil.
- Camargo, Z, Gomes, P.C., Madureira, S., Rusilo, L.C. 2015. Voice quality from a phonetic perspective: supralaryngeal and muscular tension settings. In: The Scottish Consortium for ICPhS 2015 (Ed.). Proc.18<sup>th</sup> International Congress of Phonetic Sciences, paper number 0198.1-5, Glasgow, Scotland.
- Fink, B.R., Basek, M., Epanchin, V. 1956. The mechanism of opening of the human larynx. *Laryngoscope* 66(4), 410-425.
- French, P., Foulkes, P., Harrison, P., Hughes, V., San Segundo, E., Stevens, L.C. 2015. The vocal tract as a biometric: output measures, interrelationships, and efficacy. In: The Scottish Consortium for ICPhS 2015 (Ed.). Proc.18<sup>th</sup> International Congress of Phonetic Sciences, paper number 0817.1-5, Glasgow, Scotland.
- Laver, J.1980. *The phonetic description of voice quality*. Cambridge University Press.
- Laver, J.; Mackenzie-Beck J. 2007. *Vocal Profile Analysis Scheme-VPAS*. Speech Science Research Centre, Queen Margaret University College, 2007.
- Léon P. 1993. *Précis de phonostylistique-parole et expressivité*. Paris: Nathan.
- Mackenzie-Beck, J., Schaeffler, F. 2015. Voice quality variation in Scottish adolescents: gender versus geography. In: The Scottish Consortium for ICPhS 2015 (Ed.). Proc.18<sup>th</sup> International Congress of Phonetic Sciences, paper number 0737.1-5, Glasgow, Scotland.
- Rilliard, A., d’Alessandro, C., Evrard, M. 2018. Paradigmatic variation of vowels in expressive speech: acoustic description and dimensional analysis. *The Journal of the Acoustical Society of America* 143(1), 109-122.

# On the argumenthood of optional PPs with Italian motion verbs

Giulia Cappelli<sup>1</sup>, Pier Marco Bertinetto<sup>1</sup>, Alessandro Lenci<sup>2</sup>

<sup>1</sup>Scuola Normale Superiore, Italy

<sup>2</sup>University of Pisa, Italy

DOI:

## Abstract

In this study, we investigated the argumenthood of locative PPs occurring with motion verbs in Italian, by means of a behavioral experiment based on the traditional optionality test for argumenthood. Previous research relied on linguists' intuitions or corpus frequencies, and to our knowledge no published research on motion verbs in Italian has employed a behavioral experiment yet. Consistently with recent literature, we show that the optional PPs occurring with each motion verb have different argumenthood scores on a gradient scale depending on the interaction between their Aktionsart and morphosyntactic class. We observed that achievements occur with adjuncts, activities with in-between constituents, accomplishments either with arguments or with adjuncts depending on their auxiliary in compound tenses.

Key words: argumenthood, acceptability judgments, motion verbs, Italian, syntax

## Introduction

In this paper, we are interested in the argumenthood of locative PPs occurring with motion verbs in Italian. Much has been written about it, the fragile consensus being that Goals are arguments of directed-motion verbs and other PPs are adjuncts; however, it is still common practice to rely on individual intuitions and on a binary argument/adjunct distinction.

Following Cennamo & Lenci (2018), we aim to show that differences in the argumenthood of optional PPs with motion verbs stem from the interaction between the Aktionsart of verbs and their morphosyntactic class. Here we argue for a gradient argument/adjunct distinction by means of a behavioral experiment, using PP optionality as a test for argumenthood.

## Experimental setting

### Materials and methods

Our experiment targeted 44 unique Italian motion verbs. We classified each verb according to its Aktionsart and verb class (Folli 2001, Cennamo & Lenci 2018), the latter based on auxiliary selection (see Table 1).

Table 1. Classification of motion verbs in Italian.

	Class 1 <sub>BE</sub>	Class 2 <sub>BE/HAVE</sub>	Class 3 <sub>HAVE</sub>
achievement	arrivare 'to arrive'	decollare <sub>BE/HAVE</sub> 'to take off'	-
accomplishment	cadere 'to fall'	correre <sub>BE</sub> 'to run'	-
activity	-	correre <sub>HAVE</sub> 'to run'	camminare 'to walk'

Some Class 2 activity verbs can be reclassified into accomplishments by changing their auxiliary in compound tenses from 'have' to 'be'. Under the optionality test, a given constituent is an argument if its omission results in an ungrammatical sentence, while it is an adjunct if the sentence remains grammatically correct without it. Thus, we tested the argumenthood of the PPs they occur with by using as stimuli 48 (given that 4 verbs out of 44 can have both auxiliaries) Italian sentences only consisting of a subject and a motion verb in the present perfect tense.

We recruited 40 native speakers with no training in linguistics, giving each of them 5€ as compensation. Each participant was presented with a random sample of sentences, selected so that each sentence received acceptability judgments (on a 7-point Likert scale) from 25 participants. Applying the optionality test, low acceptability scores point to the argument-like status of the optional PP, while high scores point to adjuncthood.

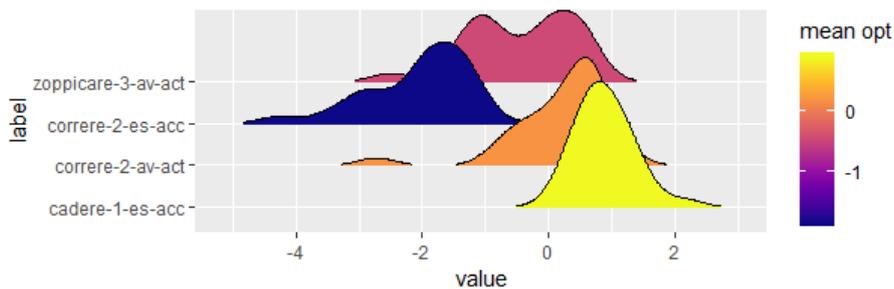


Figure 1. Density plots with z-normalized optionality scores for the verbs *zoppicare* 'to limp', *correre* 'to run', *cadere* 'to fall down'.

Following Kim *et al.* (2018), we accounted for individual differences in the use of the Likert scale by z-transforming the 25 judgments for each verb. The statistical analysis was carried out in R (R Core Team 2019), and density plots (as in Figure 1 above) were created using the R packages *ggplot2* (Wickham 2016), *ggridges* (Wilke 2018) and *viridis* (Garnier 2018).

## Results

Based on z-transformed optionality scores, there is a clear tendency for motion-verb sentences without a PP to be fully acceptable when the verbs are achievements (Mdn = .55) and Class 1 accomplishments (Mdn = .48), fully unacceptable when the verbs are Class 2 accomplishments (Mdn = -.55), halfway acceptable when the verbs are activities (Mdn = .09). Given that optionality is inversely proportional to argumenthood, we take this to mean that achievements and Class 1 accomplishments occur with adjuncts, Class 2 accomplishments with arguments, and activities with intermediate constituents (what Grimshaw 1990 called “argument-adjuncts”). A Wilcoxon rank-sum test indicated that the PPs occurring with Class 1 accomplishments are significantly more optional than the ones occurring with Class 2 accomplishments ( $W = 61716$ ,  $p < .001$ ). A Welch’s Anova test indicated that optionality scores for the three Aktionsarten are significantly different ( $F(2, 964.36) = 59.03$ ,  $p < .001$ ). Participants tend to agree more on adjuncthood than on argumenthood, based on the Pearson correlation between average score and between-subject variance ( $r(46) = -.86$ ,  $p < .001$ ).

## Conclusions

We argue that inherently telic motion verbs (i.e. achievements and Class 1 accomplishments) preferably occur with adjuncts because they have Goals as implicit arguments (Williams 2015) and their other possible PPs are adjunct-like in nature (Nam 2004). By contrast, the presence of a Goal triggers the reclassification of some activities in Class 2 into accomplishments, as shown by their shift from *have*-selecting inergatives to *be*-selecting unaccusatives (Sorace 2000 among others). Thus, Class 2 accomplishments occurring without a PP are deemed ungrammatical, because their ‘be’ auxiliary surfaces precisely to allow for the expression of an overt Goal. Now consider the intermediate status of the PPs occurring with activities. Limited to Class 2 activities, one could argue that some participants judge the PPs as adjuncts because they cannot occur with Goals, while other participants consider them as arguments because they intuitively perceive the possibility of transforming them into accomplishments, and other participants possibly provide intermediate judgments because they access both readings at once (Zubizarreta & Oh 2007, Rappaport-Hovav & Levin 2008). However, Class 3 activities cannot be reclassified into accomplishments, yet the argumenthood of PPs occurring with them is intermediate: we argue that this depends on the atelic nature of activities, which causes them to be less polarized than accomplishments and achievements with respect to argumenthood. Finally, we interpret the overall higher agreement on adjuncthood than on argumenthood as proof that motion verbs tend indeed

not to select for arguments: thus, participants tend to preferably activate the adjunct reading even when faced with an argument-selecting verb.

## Acknowledgements

We thank Najoung Kim for prompting this study and for her invaluable advice, and the anonymous reviewers for their sharp critique of our abstract.

## References

- Cennamo, M., Lenci, A. 2018. Gradience in Subcategorization? Locative Phrases with Italian Verbs of Motion. *Studia Linguistica*, 1–29.
- Folli, R. 2001. Two strategies to construct Telicity: A comparative analysis of English and Italian. *Working Papers in Linguistics, Philology & Phonetics*, 47.
- Garnier, S. 2018. *viridis*: Default Color Maps from 'matplotlib'. R package version 0.5.1. <https://CRAN.R-project.org/package=viridis>
- Grimshaw, J. 1990. *Argument structure*. The MIT Press.
- Kim, N., Rawlins, K., Van Durme, B., Smolensky, P. 2018. Predicting the Argumenthood of English Prepositional Phrases. Retrieved from <http://arxiv.org/abs/1809.07889>
- Levin, B., Rappaport Hovav, M. 2008. Lexicalized manner and result are in complementary distribution. *IATL*, 26–27.
- Nam, S. 2004. Goal and Source: Their Syntactic and Semantic Asymmetry. *Annual Meeting of the Berkeley Linguistics Society*, 301.
- R Core Team 2019. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Sorace, A. 2000. Gradients in auxiliary selection with intransitive verbs. *Language*, 859–890.
- Vater, H. 1978. On the possibility of distinguishing between complements and adjuncts. *Valence, Semantic Case and Grammatical Relations*, 1, 21–45.
- Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- Wilke, C. O. 2018. *ggridges: Ridgeline Plots in 'ggplot2'*. R package version 0.5.1. <https://CRAN.R-project.org/package=ggridges>
- Williams, A. 2015. *Arguments in Syntax and Semantics*.
- Zubizarreta, M. L., Oh, E. 2007. *On the syntactic composition of manner and motion*. The MIT Press.

# Russian homophones and visual word recognition

Daria Chernova

Laboratory for Cognitive Studies, Saint Petersburg State University, Russia

DOI:

## Abstract

Our study focuses on visual word processing in silent reading for comprehension. We compare the processing of homophone errors and non-homophonous orthographic neighbours with processing of target words in congruent sentential context. We find that homophone errors are the most costly in processing as they provoke the competition between the phonological and visual representation of the word.

Key words: silent reading, phonological representation, orthographic representation, homophones

## Introduction

The role of phonology in visual word recognition is a crucial theoretical issue in psycholinguistics research of reading (Clifton 2015).

The main question is what is stored in mental lexicon – phonological representations mapped onto semantics or both phonological and orthographic ones?

The dual-route model (Coltheart 1980) argues for two independent ways of printed words processing; one relies on phonological representation and the other on the visual image of the word. For frequent words processed by skilled readers, the visual route to the meaning is accessed faster (Jared et al. 1999).

The single-route theory (Frost 1998) claims that phonological processing is a default procedure of cognitive system and meaning only can be accessed via phonological representation. Strong phonological models claim phonological processing to be mandatory and automatized.

Homophones can be a suitable material for testing these theories (Van Orden 1987, Newman et al. 2012): if a homophone is processed in the same way as any other orthographic neighbour, it gives evidence for the visual route to meaning. If homophony facilitates processing, it gives evidence for the phonological route.

## Materials and design

Most studies on visual processing of homophones have been carried out on English. We address to Russian language, which has more transparent orthography and regular homophony - for example, words with voiced and

devoiced consonants in the end (e.g. *kot* /kot/ ‘cat’- *kod* /kot/ ‘code’), which differ by one letter as other orthographic neighbours do.

We constructed three types of stimuli: a correct target word congruent with the context (a), a target word substituted with a homophone (b), and a target word substituted with a non-homophonous orthographic neighbor, i.e. a spelling control (c).

- a. *Pushistyj kot svernulsia klubkom v kresle.*  
‘The fluffy cat curled up in the armchair’.
- b. *Pushistyj kod svernulsia klubkom v kresle .*  
‘The fluffy code curled up in the armchair’.
- c. *Pushistyj kol svernulsia klubkom v kresle*  
‘The fluffy stake curled up in the armchair’.

All sentences had 6 words and the target word was always the second one. The materials consisted of 24 sets of target sentences (8 sentences for each condition) and 60 filler sentences which contained no errors. The sentences were distributed into three experimental lists according to Latin square design.

## Method and participants

39 native speakers of Russian aged 18-32 performed a word-by-word noncumulative self-paced-reading task and answered comprehension questions after 30% of sentences. The experiment was run online using Ibx Farm platform.

## Results

We analyzed participants’ question-answering accuracy and reading times. The raw reading times (per word) that exceeded 1500 ms were adjusted to this threshold (about 3% of the data). As for question-answering accuracy, no participant made more than 5 mistakes, so all subjects were included in the analysis.

No significant differences were found in the processing time of the target word (region 2) in all of the conditions. However, on the word which follows the target one (region 3) t-test shows significant difference between conditions: reaction time for the non-homophonous orthographic neighbour is significantly longer in comparison to the correct word ( $t=3.5$ ,  $p=0.002$ ), whereas the reaction time for the homophone is the longer in comparison to the orthographic neighbor ( $t=2.9$ ,  $p=0.001$ ) as well as to the control condition ( $t=3.1$ ,  $p=0.0009$ ). Thus, homophone is the

most costly in processing despite the fact its phonological representation coincides with the contextually consistent word. These differences reveal as spill-over effects as soon as they have to do with contextual integration process.

As for the wrap-up effect, the last word in correct sentences is processed faster in comparison to the last word in sentences containing both types of errors.

The time course of sentence processing in three conditions is shown in Figure 1.

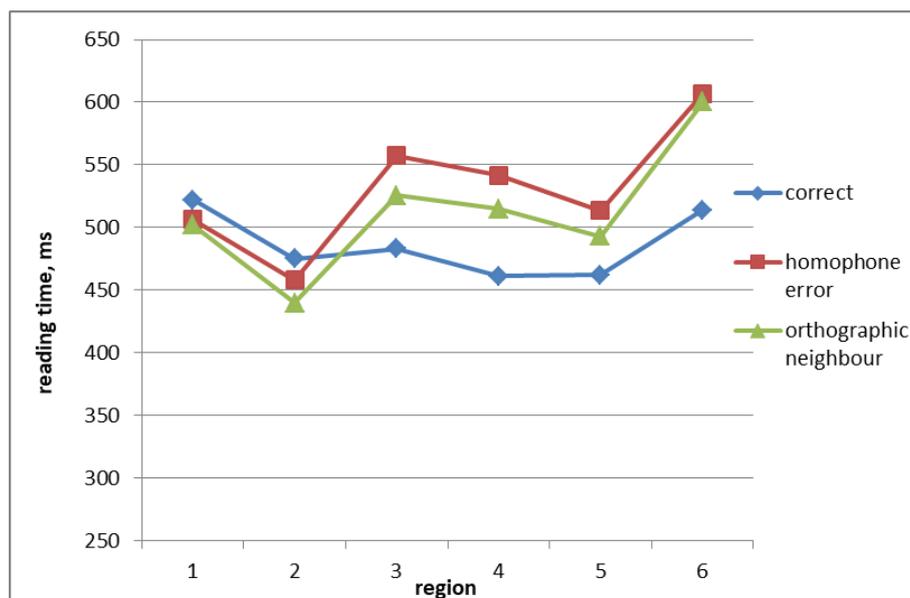


Figure 1. Time course of sentence processing in three conditions

## Discussion

Our data is compatible with dual-route model. Orthographic neighbour instead of the target word slows down the processing due to context inconsistency. As for the homophone, it provokes the competition between phonological and visual representation of the word. Phonological representation activates both the consistent and the inconsistent meaning while visual representation activates the inconsistent one. The competition between the meanings makes the processing even more costly than processing of the inconsistent orthographic neighbor. Thus, homophony rather impedes than facilitates visual word processing.

## **Acknowledgements**

The reported study was funded by RFBR according to the research project № 18-00-0646 (18-00-00640).

## **References**

- Clifton, C. 2015. The Roles of Phonology in Silent Reading: A Selective Review. In: Frazier L., Gibson E. (eds) *Explicit and Implicit Prosody in Sentence Processing. Studies in Theoretical Psycholinguistics*, vol 46. Springer.
- Coltheart, M. 1980. Reading phonological receding and deep dyslexia. In M. Coltheart, K. Patterson, J.C. Marshal (eds.), *Deep dyslexia* (pp. 197-226). London: Routledge & K. Paul.
- Frost, R. 1998. Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, 123, 71–99.
- Jared, D., O'Donnell, K. 2017. Skilled adult readers activate the meanings of high frequency words using phonology: Evidence from eye tracking. *Memory & Cognition*. 43. 334-346.
- Newman R.L., Jared D., Haigh C.A. 2012. Does phonology play a role when skilled readers read high-frequency words? Evidence from ERPs. *Language and Cognitive Processes*, 27:9, 1361-1384.
- Van Orden, G.C. 1987. A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, 15, 181-198.

# Image effects and long distance dependencies

Leticia M.S. Corrêa<sup>1</sup>, Marina R.A. Augusto<sup>2</sup>

<sup>1</sup>Department of Letters, Pontifical Catholic University of Rio de Janeiro, Brazil

<sup>2</sup>Department of Language Studies, Rio de Janeiro State University, Brazil

DOI:

## Abstract

This study aimed to distinguish the syntactic effect of syntactic movement/featural intervention from the post-syntactic effect of image complexity in children's comprehension of object relative clauses and WH-questions in a picture-identification task. Data from Brazilian Portuguese-speaking seven-year-olds are analysed. The results indicate that image complexity amplifies the number of errors that might be ascribed to an intervention effect. Similar results on WH-questions suggest that the effect of syntactic movement at the interfaces suffices to account for the subject-object asymmetry. Implications for the identification of language impairment in the syntactic domain are considered.

Key words: relative clauses, sentence comprehension, featural intervention

## Introduction

The subject-object asymmetry in children's comprehension of relative clauses (RCs) with object RCs (ORCs) being more demanding than subject RCs (SRCs) has been widely attested (Adani 2011, Corrêa 1995, de Villiers, Hakuta, Cohen 1979, Diessel, Tomasello 2000) (1-2).

- (1) Show me the horse that kicked the cow. (SRC)
- (2) Show me the horse that the cow kicked. (ORC)

A number of explanations for this asymmetry have been provided, focusing on the possible consequences of syntactic movement on the processing load (Corrêa 1995, Gibson, Desmet, Grodner, Watson, Ko 2005, Wanner, Maratsos 1978). In a different vein, the featural intervention hypothesis proposes that the subject of ORCs functions as an intervener thereby blocking syntactic movement, if the Relativized Minimality principle (Rizzi 1990) is overextended or generalized (Friedmann, Belletti, Rizzi 2009, Friedmann, Novogrodsky 2006, Grillo 2009). The relevant feature distinguishing the subject and the object of the RC, thereby promoting syntactic movement (WH+move) would fail to be recognized, retrieved or maintained in working memory.

The assessment of children's syntactic abilities has been typically conducted in the sentence-picture matching paradigm, which contributes to

the identification of language impairment in the syntactic domain. Post-syntactic processes pertaining to the mapping of the semantic interpretation of the sentence onto pictures may, nevertheless, add to the overall processing difficulty. The aim of this study is to tease apart the effect of syntactic movement/featural intervention from post-syntactic effects stemming from image complexity on the comprehension of RCs by children.

### Syntactic demands and image complexity

In the present study, a battery of tests for the assessment of language comprehension abilities of Brazilian Portuguese (BP) schoolchildren (MABILIN, Correa, 2000) was used, in which *image complexity* was manipulated. Data obtained in the comprehension of restrictive right-branching RCs (3-4) and bare WH-questions (5-6) are analysed here. RCs would be vulnerable to featural intervention, unlike bare WH-questions. For the latter, subject-object asymmetry would more likely be due to the consequences of syntactic movement to sentence processing.

- (3) Mostra o urso que pegou o tigre. (SRC)  
Show the bear that caught the tiger.
- (4) *Mostra o leão que o tigre pegou.* (ORC)  
Show the lion that the tiger caught.
- (5) Quem limpou o elefante? (Bare SWH-Q)  
Who cleaned the elephant?
- (6) Quem o porco chamou? (Bare OWH-Q)  
Who did the pig call?

For each test-sentence, a background picture is provided with the relevant event introducing the referent of the +WH element as the actor/ patient of the action (the target referent). Children are requested to identify this referent in a set of three options. The complexity of the images is manipulated as follows: Simple images (SI) present a single event (the relevant event) and a third character standing by the scene. Complex images (CI) present two similar events: the target one and another one with reversed roles. (See Figures 1-3 for SI and 2-4 for CI ORCs and OWH-Qs).

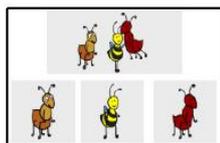


Fig.1 SI ORC

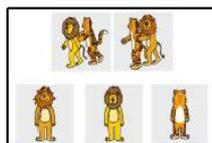


Fig.2.CI ORC

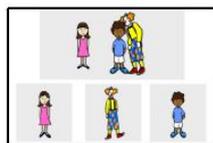


Fig.3 SI OWH-Q

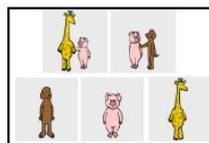


Fig.4 CI OWH-Q

For both sentence types, the independent variables were *image complexity* (simple and complex) and *moved element* (subject and object). The dependent variables were the number of target responses; the number of referential errors (the second token of the same type as the target one in RCs/the additional character in WH-questions) and the number of subject interference/intervention errors in object structures.

It was expected that the subject-object asymmetry was more evident in RCs. Complex images were predicted to be more demanding. This demand was expected to affect ORCs in particular and be manifested in the subject interference/intervention errors.

The participants were sixty-eight 7-year-olds (37 girls, mean age 7.5) Brazilian Portuguese (BP) speakers, enrolled in the Rio de Janeiro public school system (predominantly low-middle social-economic status), with no report of language or other cognitive impairment. The material consisted of eight sentences in each condition defined by *moved element* presented either in a simple or complex image condition. There were, therefore, four trials for each of the test conditions. A laptop computer and the MABILIN program were used. Children were invited to play a computer game with the experimenter in an isolated room of the school. The sentences were orally presented by the experimenter. The test took about 25 minutes.

## Results

For both RCs and WH-questions, correct responses (1) were submitted to a 2x2 ANOVA with two within-subject factors. Referential (2) and subject interference/intervention errors (3) were analysed by means of the Friedman test for repeated measures.

For RCs: (1): main effects of *image* ( $F(1,67)=26.2$   $p<.001$ ) (Means: 3.76 (SI); 3.37 (CI)) and *moved element* ( $F(1,67)=24.5$   $p<.001$ ) (Means: 3.75 (S); 3.38 (O)); 2-way interaction ( $F(1,67)=13.1$   $p<.001$ ) (Means: 3.81(SI S); 3.71(SI O); 3.69 (CI S); 3.04 (CI O)). (2): No significant effects. (3): *image* effect:  $X^2r = 12.3676$  (1, N = 68),  $p <.0001$  (Sums: 87.5 (SI); 116.5 (CI)).

For WH-questions: (1) main effects of *image* ( $F(1,67) = 10.5$   $p<.01$ ) (Means: 3.9 (SI); 3,7 (CI)) and *moved element* ( $F(1,67)=12.4$   $p<.001$ ) (Means: 3.89 (S); 3,71 (O)). (2): No significant effects. (3): *image* effect:  $X^2r = 5.3088$  (1, N = 68),  $p = .02$  (Sums : 92.5 (SI); 111.5 (CI)).

## Discussion and final remarks

The results show that for RCs, a clear subject-object asymmetry can be specially observed in the complex image condition and the demands of the object structures are particularly manifested in the subject interference/intervention errors. That is, children are more likely to point to the referent of the intervening subject. For WH-questions, there is no

interaction between image and moved element in the number of correct responses, suggesting that the impact of the complex image does not affect the successful processing of these sentences, even though subject interference/intervention errors are also more likely to occur in the complex image condition. The predictions are thus confirmed. Image complexity amplifies the number of errors that might be ascribed to an intervention effect in ORCs or to subject interference in both types of sentences, though affecting RCs in particular. The fact that these errors are post-syntactic and were also amplified in object WH-questions causes into question the intervention hypothesis. It appears that the effect of syntactic movement on sentence processing at the interfaces suffices to account for the subject-object asymmetry, regardless of RM. These results should be taken into account when the assessment of language abilities is considered for the identification of language impairment in children. Even though the processing of long-distance dependencies in the most demanding conditions can distinguish linguistic abilities, it is the difficulty imposed by object structures with simple images that would more likely provide an alarm sign for language impairment.

### **Acknowledgements**

FAPERJ: E-24/202.959/2017;CAPES-PRINT; Prociência FAPERJ.

### **References**

- Adani, F. 2011. Rethinking the acquisition of relative clauses in Italian: Towards a grammatically based account. *Journal of Child Language*, 38(1), 141–165.
- Corrêa, L.M.S. 1995. The relative difficulty of children's comprehension of relative clauses: A procedural account. In Nelson, K.E., Réger, Z. (Ed.), *Children's Language* (vol.8) (pp. 225–244). Hillsdale: Lawrence Erlbaum.
- de Villiers, J.G., Hakuta, K., Cohen, M. 1979. Children's comprehension of relative clauses. *Journal of Psycholinguistic Research*, 8(5), 499–518.
- Diessel, H., Tomasello, M. 2000. The development of relative clauses in spontaneous child speech. *Cognitive Linguistics*, 11(1/2), 131–151.
- Friedmann, N., Belletti, A., Rizzi, L. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, 119(1), 67–88.
- Gibson, E., Desmet, T., Grodner, D., Watson, D., Ko, K. 2005. Reading relative clauses in English. *Cognitive Linguistics*, 16(2), 313–353.
- Grillo, N. 2009. Generalized Minimality: Feature impoverishment and comprehension deficits in agrammatism. *Lingua*, 119(10), 1426–1443.
- Wanner, E., Maratsos, M. 1978. An ATN approach to comprehension. In Halle, M, Bresnan, J., Miller, G.A. (Ed.), *Linguistic Theory and Psychological Reality* (pp. 119–161). Cambridge, MA: MIT Press.

# On the processing of object relative clauses

Letícia M.S. Corrêa<sup>1</sup>, Erica dos S. Rodrigues<sup>1</sup>, Renê Forster<sup>2</sup>

<sup>1</sup>Department of Letters, Pontifical Catholic University of Rio de Janeiro, Brazil

<sup>2</sup>Department of Language Studies, Rio de Janeiro State University

DOI:

## Abstract

This study investigates the effect of featural mismatch in the processing of object relative clauses (ORCs). A self-paced reading experiment is reported in which the matching of formal and semantic number features of the head noun and the RC subject is manipulated. The predicted full mismatching intervention/interference effect was not obtained but collective nouns gave rise to the shortest reading times in both the subject and the RC verb regions. A follow-up suggests that it is subject-verb agreement in both the RC and the main clause that is vulnerable to semantic interference.

Key words: relative clauses, featural intervention, interference

## Introduction

The asymmetry between subject and object relative clauses (RCs) is well-documented (Gibson, Desmet, Grodner, Watson, Ko, 2005; Wanner, Maratsos, 1978). Object RCs (ORCs) impose greater demands than subject RCs (SRCs). These demands can be nevertheless minimized in particular conditions. It has been demonstrated that dissimilarity between the RC head (N1) and subject (N2) can facilitate processing (Gordon, Hendrick, Johnson, 2004; Gordon, Hendrick, Levine, 2002). Mismatching effects can be due to encoding interference (Villata, Tabor, Franck, 2018), to retrieval interference, in a cue-based approach (Van Dyke, McElree, 2006) or to both effects (Villata et al., 2018). From a linguistic perspective, structural dissimilarity could be interpreted as reflecting featural intervention (Adani, 2011; Grillo, 2009; Villata, Rizzi, Franck, 2016), in the light of the Relativized Minimality principle, which blocks object movement under identity of features (Rizzi, 1990). The more similar N1 and N2 are, the more likely it is for this principle to apply.

The present study investigates the featural mismatching effect in the processing of ORCs by manipulating the syntactic and semantic number features of N1 and N2 in a self-paced reading comprehension task. Featural intervention would predict faster reading times in the plural condition and similar times for singular and collective nouns, since it applies to formal features only. Memory interference might allow semantic effects and would predict a gradient from full to partial mismatching. Two experiments were

conducted with adult native speakers of Portuguese, in which the number features of N2 and N1 were manipulated, respectively.

In Experiment 1 (n=42, 34 females; mean age 24 years old), the critical stimuli were distributed in three conditions: full matching (singular) (1); full mismatching (plural) (2); partial mismatching (collective) (3). 3 stimuli per condition and 18 fillers were presented, in a Latin square design, followed by a comprehension question. The critical regions were the RC subject (N2) (region 1) and the RC verb (region 2). Both raw and residual reading times were computed for each region. The task was conducted on a Windows based laptop running the software Linger.

- singular-N2: *O empresário que [o jogador REGION 1] de futebol*  
 (1) *[contratou REGION 2] aposta na renovação da equipe.* (The manager that the player of soccer hired bets in the renewal of the team.)  
 (plural-N2: *O empresário que [os jogadores REGION 1] de futebol*  
 2) *[contrataram REGION 2] ...* (The manager that the players of soccer hired ...)  
 ( collective- N2: *O empresário que [o time REGION 1] de futebol*  
 3) *[contratou REGION 2] ...* (The manager that the team of soccer hired...)

For Experiment 2 (n=30, 26 females; mean age 22), design and procedure were similar to those of experiment 1, except for the independent variable (number) on the head noun (N1). The 18 critical sentences were distributed in the same 3 conditions of experiment 1 (4-6) with 36 fillers. The main clause verb region (region 3) was added. Residuals were computed for this region only.

- singular-N1: *O jogador que [o empresário REGION 1] [contratou*  
 (4) *REGION 2] [aposta REGION 3] na renovação da equipe.* (The player that the manager hired bets in the renewal of the team.)  
 (plural-N1: *Os jogadores que [o empresário REGION 1] [contratou*  
 5) *REGION 2] [apostam REGION 3] ...* (The players that the manager hired bet ....)  
 ( collective-N1: *O time que [o empresário REGION 1] [contratou*  
 6) *REGION 2] [aposta REGION 3] ...* (The team that the manager hired bet....).

## Results

The data were analysed by means of a three-way ANOVA in which *number* was a within-subject factor (see Table 1 for the mean reading times).

In experiment 1, the effect of number was significant in both regions and the direction of the means was similar for the raw and the residual measures. Raw on N2 ( $F(2,82) = 5.10$   $p < .01$ ; on RCV ( $F(2,82) = 3.58$   $p = .03$ ; residual on N2 ( $F(2,82) = 5.31$   $p < .01$ ); on RCV ( $F(2,82) = 6.40$   $p < .01$ ). This effect was not in the predicted direction: Full mismatching (plural) was not significantly different from the full matching (singular) in any analysis. The shortest reading times were obtained with collective nouns (partial mismatch), which rules out a featural intervention effect on RCV. The difference between collective and plural was significant in all analyses but a gradient with faster times in the plural condition was not obtained. The effect captured on the collective condition on N2 is compatible with encoding interference, due to the mismatch of the syntactic/semantic number features of N1 and N2. The effect on the RC verb is compatible with retrieval interference, if a collective noun affects the reactivation of a partially feature sharing object, with encoding interference and with an effect on subject-verb agreement within the RC. Subject-verb agreement can be vulnerable to collective words (Kreiner, Garrod & Sturt, 2013) and a semantic effect has been captured in ORCs (Fedorenko, Gibson & Rohde, 2006).

Conditions	Experiment 1 (N1 sing.)				Experiment 2 (N2 sing.)			
	N2 raw**	res.**	RCV raw*	res.**	N2 raw †	RCV raw	MCV raw††	res††
N2 sing.	1004	- 1.5	1194	195				
N2 pl.	1115	- 88.7	1202	152				
N2 coll.	885	- 228	989	-96.2				
N1 sing.					728	1026	890	150
N1pl.					808	929	893	74
N1 coll.					725	934	770	24

Table 1. Mean reading times (msec) per condition and region in each experiment (\*Effect of number: \*  $p < .05$ ; \*\*  $p < .01$ ; †  $p = .05$ ; ††  $p < .1$ )

Experiment 2 was a follow-up intended to clarify this point. The pattern of the results of Experiment 1 was maintained. The marginally significant effect on N2 was compatible with an encoding interference effect of collective nouns. Retrieval interference in the reactivation of the moved object would be expected to affect the RC verb due to a plural feature in N1 (Wagers, Lau & Phillips, 2009). It was, however, the main clause verb that was weakly affected (particularly due to sing. vs coll. on residual responses ( $t(29)=2.28$   $p = .03$ ). Even though a spill-over effect of the RCV cannot be ruled out, the effect on MCV supports the view that it is the subject-verb agreement by itself that is vulnerable to partial number mismatching.

## Final remarks

The present results suggest that Portuguese-speaking adults are not likely to be affected by number intervention or retrieval interference in the syntactic processing of ORCs, at least when no additional material is included between the RC subject and verb. The semantic effect obtained does not seem to bear on the actual processing of the object displacement in the RC. It is more likely to apply to subject-verb agreement. Both the embedded clause and the main clause are vulnerable to the encoding of a collective subject, which affects subject-verb number agreement.

## Acknowledgements

CNQq PQ-309499/2016-0 and CAPES-PRINT 88881.310705/2018-01

## References

- Adani, F. 2011. Rethinking the acquisition of relative clauses in Italian: Towards a grammatically based account. *Journal of Child Language*, 38(1), 141–165.
- Fedorenko, E., Gibson, E., Rohde, D. 2006. The nature of working memory capacity in sentence comprehension. *Journal of Memory and Language*, 54(4), 541–553.
- Gibson, E., Desmet, T., Grodner, D., Watson, D., Ko, K. 2005. Reading relative clauses in English. *Cognitive Linguistics*, 16(2).
- Gordon, P. C., Hendrick, R., Johnson, M. 2004. Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, 51(1), 97–114.
- Gordon, P.C., Hendrick, R., Levine, W.H. 2002. Memory load interference in syntactic processing. *Psychological Science* 13 (5), 425–30.
- Grillo, N. 2009. Generalized Minimality: Feature impoverishment and comprehension deficits in agrammatism. *Lingua*, 119(10), 1426–1443.
- Kreiner, H., Garrod, S., Sturt, P. 2013. Number agreement in sentence comprehension. *Language and Cognitive Processes*, 28(6), 829–874.
- Rizzi, L. 1990. *Relativized Minimality* (Linguistic Inquiry Monographs 16). Cambridge, MA: MIT Press.
- Van Dyke, J. A., McElree, B. 2006. Retrieval interference in sentence comprehension. *Journal of Memory and Language*, 55(2), 157–166.
- Villata, S., Rizzi, L., Franck, J. 2016. Intervention effects and Relativized Minimality: New experimental evidence from graded judgments. *Lingua*, 179,
- Villata, S., Tabor, W., Franck, J. 2018. Encoding and Retrieval Interference in Sentence Comprehension. *Frontiers in Psychology*, 9 (2), 1-16.
- Wagers, M., Lau, E. Phillips. C. 2009. Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61 (2), 206-237.
- Wanner, E., Maratsos, M. 1978. An ATN approach to comprehension. In Halle, G. A., Morris, A., Bresnan, J. Miller (Eds.), *Linguistic Theory and Psychological Reality* (pp. 119–161). Cambridge, MA: MIT Press.

# Automatic administration of semantic verbal fluency tests for Portuguese

João Costa, Ana Pereira, Liliana Ferreira

Fraunhofer Portugal AICOS

DOI:

## Abstract

Verbal fluency tests are quick and flexible tests used in the area of Neuropsychology to evaluate executive and language functions of subjects. An automatic modular pipeline for the analysis of verbal fluency tests for the semantic category of animals is proposed in this document. The system was developed for European Portuguese using test audio recordings as input. Google Cloud Speech-to-Text is used to perform automatic speech recognition. *fastText* word embeddings and phonemic transcriptions are applied to automatically evaluate temporal, phonemic, and semantic clusters produced. The proposed architecture was evaluated using 164 animal category tests, performed by Portuguese elderly subjects. Correlations were found between manual and automatic extracted features, such as the number of correct words produced ( $p=0.58$ ). Feature extraction efficacy was found to be dependent on background noise levels during recording.

Key words: verbal fluency tests, natural language processing, automatic speech recognition.

## Introduction

Verbal Fluency Tests (VFT) are frequently used to assess verbal and executive functions of subjects, as a standalone cognitive test or as part of a neuropsychological test battery. These tests consist in the production of as many words as possible, with respect to a given rule. For semantic VFTs, subjects are asked to produce words that belong to the same semantic category within a one minute time window; animal names are frequently used for this task.

In this work, a fully automatic and modular architecture for VFT administration is proposed. The developed pipeline was constructed for use with semantic VFT (animal names) in European Portuguese. With the proposed system, recorded audio from a VFT can be automatically transcribed, filtered, and processed, thus quickening and improving analysis of these tests, with the extraction of clinically relevant features.

## Automatic administration of verbal fluency tests

The pipeline of the proposed system for automatic administration and analysis of VF tests is shown in Figure 1. The system was implemented in

Python 3.6, and uses as input audio recordings of VFT execution. The developed solution can extract 20 metrics from a single test.

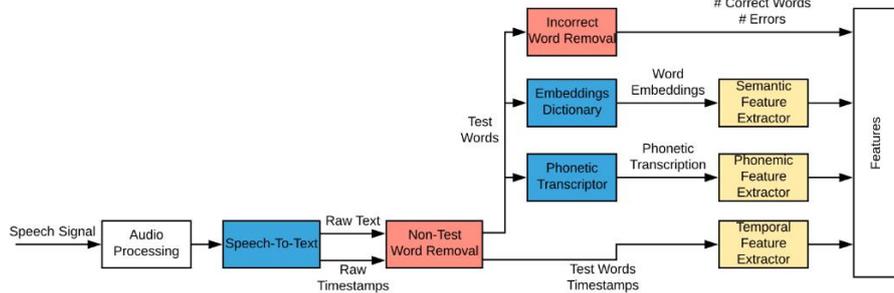


Figure 1. Overview of developed pipeline for automatic transcription and analysis of VFTs.

### Speech-to-text

To perform Automatic Speech Recognition (ASR), the Google Cloud Speech-to-Text (Google Cloud 2019) service is used. This module converts each audio recording to a list of transcribed words, in European Portuguese, with the corresponding word timestamps in the audio signal.

### Feature Extraction

Using the test transcript obtained after word filtering, a total of 20 features are extracted for each VF test, namely, the number of correct words produced, the number of errors, six temporal features, six semantic features, and six phonemic features.

Temporal features are derived from the predicted temporal clusters, defined using the DBSCAN clustering algorithm (Xiaowei 2002), from the word timestamps provided by Google Cloud Speech-to-Text.

Phonemic features are derived from the predicted phonemic clusters, which are based on the phonemic similarity between VFT exemplars. The phonemic similarity of a word pair is calculated using the Levenshtein distance between their phonetic transcriptions, provided by Infopédia (Porto Editora 2019).

Semantic features are derived from the predicted semantic clusters, which are based on the semantic similarity between VFT exemplars. The semantic similarity of a word pair is calculated using the cosine distance between their corresponding *fastText* deep word embeddings (Grave 2018).

### Data collection

To validate the developed solution, 164 semantic VFT (animal names) recordings were collected with a *smartphone* from Portuguese elderly subjects. Recordings were performed in the scope of the clinical trial

performed by Amorim & Susano (2017). Manual transcription was performed by the first author, annotating all test exemplars (including errors, repetitions, and word variations).

This study was approved by the competent bodies *Departamento de Ensino Formação e Investigação*, *Comissão de Ética*, and *Conselho de Administração* of Centro Hospitalar do Porto, approval number 253/16 (216-DEFI/ 205-CES).

## Results

### ASR Performance

The Verbal Fluency Word Error Rate (VFWER) (König 2018) is used to evaluate the performance of the selected Speech-to-Text system (Figure 2). Performance of automatic transcription of VFTs was found to be highly dependent on background noise levels present in recordings, due to competing sounds and voices that obfuscate the speech of the subject performing the VFT.

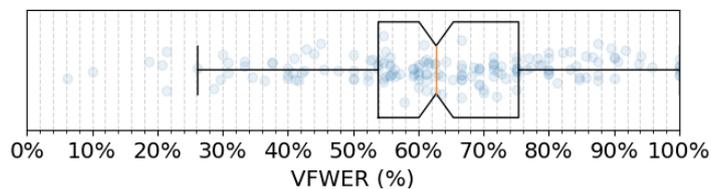


Figure 2. Verbal Fluency Word Error Rate distribution.

### Feature Extraction

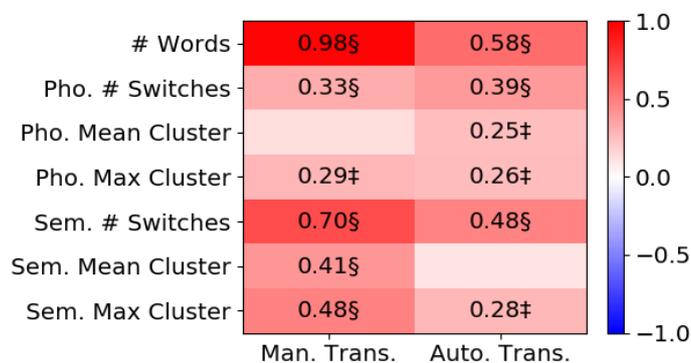


Figure 3.  $\rho$  of analogous feature pairs for manual and automatic transcriptions of VFTs ( $‡p < 0.001$ ;  $§p < 0.0001$ ).

Spearman's  $\rho$  was used to determine correlation between manually and automatically determined analogous feature pairs, using both manual and automatic transcriptions of VFTs (Figure 3). The high levels of VFWER lead to weaker correlations for automatic transcripts of VFTs. Different approaches used for manual and automatic clustering may contribute for lower correlations, with the former based on strict rule sets, and the latter on more generalised automatic approaches by word embeddings and phonetic transcript similarities.

## Conclusions

The new method of VFT analysis proposed in this document can potentially be used in the future in clinical settings by health professionals. Since extracted metrics have predictive power and can be used as screening tests for neurocognitive disorders, automatic VFTs can become a very powerful, easy to implement, and quick to apply tool to aid screening and diagnosis. Additionally, the modular design used for the implementation of the tool facilitates its adaptation to other tasks and languages.

## Acknowledgements

The authors would like to thank Dr. Pedro Amorim, Dr. Maria João Susano, Dr. Francisco Seixas and Mariana Dias of Centro Hospitalar do Porto, and Prof. Aníbal Ferreira of Faculdade de Engenharia da Universidade do Porto for their interest, availability, and data sharing.

## References

- Amorim, P., Susano, M. J. 2017. Clinical Trial: Pre-operative Neurocognitive Disorder and Low Near-infrared Spectrometry is Associated with Postoperative Delirium. <https://clinicaltrials.gov/ct2/show/NCT03171766?id=NCT03171766>
- Google Cloud. 2019. Cloud Speech-to-Text - Speech Recognition | Cloud Speech-to-Text API. <https://cloud.google.com/speech-to-text/>
- Grave, E. et al. 2018. Learning Word Vectors for 157 Languages.
- König, A. et al. 2018. Fully automatic speech-based analysis of the semantic verbal fluency task. *Dementia and Geriatric Cognitive Disorders*, 45(3-4):198–209.
- Porto Editora. 2019. Infopédia - Dicionários Porto Editora. <https://www.infopedia.pt/>
- Troyer, A.K. et al. 1997. Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, 11(1), 138–146.
- Xiaowei Xu et al. 2002. A distribution-based clustering algorithm for mining in large spatial databases. *Proceedings 14th International Conference on Data Engineering*, 324–331.

# Auditory priming of pseudo-suffixed words

Ava Creemers, Nattanun Chanchaochai, David Embick  
Department of Linguistics, University of Pennsylvania, USA  
DOI:

## Abstract

We present results from two auditory experiments that examine priming effects for pseudo-derived pairs of words (e.g., *corner-corn*), as compared to morphologically, phonologically, and semantically related pairs. Previous work shows facilitation for pseudo-derived pairs in a visual masked priming paradigm; our Experiment 1 presents the novel finding that these effects are also found in auditory priming, with no consistent difference between pseudo-derived and phonologically-related pairs. Experiment 2 introduces an intervener between primes and targets to further probe the nature of the facilitation in pseudo-derived pairs. Implications for theories of morphological processing and spoken-word recognition are discussed.

Key words: auditory word recognition, morphological priming, pseudo-suffixation

## Introduction

Visual masked priming experiments have focused on different relations between words and their embedded parts, including prime-target pairs that are morphologically and semantically related (*teacher-teach*), and pseudo-derived pairs like *corner-corn*. In the latter, *corner* occurs with an existing stem and suffix but is not etymologically or semantically related to *corn*. Pseudo-derived words have been used to argue for orthography-driven decomposition of words (Rastle et al. 2004, Rastle and Davis 2008, i.a.), for morphological decomposition independent of semantic transparency (Marslen-Wilson et al. 2008, i.a.), and for both sides in the debate between form-then-meaning (Beyersmann et al. 2016, Longtin et al. 2003) and form-and-meaning (Feldman et al. 2009; 2015, Milin et al. 2017) approaches.

Some questions are unanswered: Are decomposition effects for pseudo-derived items restricted to orthography (cf. Rastle, Davis 2008), or can they also be found in the auditory modality? And, are effects for pseudo-derived words, which do not actually share a morphological relation with their stem, morphological in nature, or merely form-driven? To investigate these issues, we present two primed lexical decision experiments with auditorily presented primes and targets. Experiment 1 shows that priming effects for pseudo-derived words can be obtained in the auditory modality. Experiment 2 investigates the nature of these effects by including an intervening item between primes and targets, based on the idea that phonological facilitation decays more rapidly than morphological facilitation.

## Experiments 1 and 2

### Methods

Conditions include prime–target pairs that are morphologically and semantically related (MS: *creamy-cream*, *treatment-treat*), pseudo-morphologically related ((pseudo-)M: *belly-bell*, *pigment-pig*), purely phonologically related with no (pseudo-)suffix (Ph: *dogma-dog*, *pillow-pill*), and semantically related (S: *garbage-trash*, *painting-art*). The pseudo-M and MS conditions differ as minimally as possible, with different suffixes appearing equally often in these conditions. M pairs that share a remote or archaic meaning (e.g., *archer*) were excluded. Targets are orthographic and phonological sub-strings of their prime. The meaning of the prime in MS pairs can always be derived from the meaning of the stem, as evidenced by high pairwise Latent Semantic Analysis (LSA) measures. S primes were also selected based on high LSA measures with the target. Targets are monosyllabic; primes are disyllabic with main stress on the first syllable.

In total, we included 20 prime-target pairs per condition. Unrelated primes were selected for each pair. Subject saw 80 critical pairs of which half were unrelated. Subjects were allocated to one of two lists, so that they saw each target only once. For the purpose of the lexical decision task, 100 filler words and 260 non-words were included. Finally, to investigate the time-course with which different types of linguistic information become available, we manipulated the Inter-Stimulus Interval (ISI) as a between-subject factor in Experiment 1. We included ISIs of 200-400ms (short-ISI), 600-800ms (medium-ISI), and 1000-12000ms (long-ISI). In Experiment 2, we used a lag of 1 intervening item and a medium-ISI between prime and target to investigate the nature of the effects in the M condition, aiming to tease apart phonological and morphological effects.

Stimuli were recorded by an adult native speaker of American English. A lexical decision task was implemented in PsychoPy2. Participants were native speakers of English (122 in Experiment 1; 80 in Experiment 2), who received course credit as compensation for participation. They were instructed that they would hear existing and non-existing words and had to make a lexical decision to each word as quickly and accurately as possible.

### Results Experiment 1

Responses were coded for response type and response time (RT). Differences in duration of the sound files were included in the model. Incorrect responses to primes and targets were discarded, and one subject was excluded due to low accuracy. We combine a-priori data trimming with model criticism, and analyze log-transformed RTs to targets with linear mixed-effects models.

The results are illustrated in Figure 1. Collapsing the different ISIs used, significant priming effects are found in all conditions ( $p < 0.001$  for MS, M, Ph, S), which also holds for each individual ISI-version. The model indicates a larger effect in the MS condition compared to the M ( $p < 0.001$ ), Ph ( $p < 0.001$ ), and S ( $p < 0.001$ ) conditions, which again holds regardless of ISI-group. No difference is found between M and Ph ( $p = 0.956$ ) in the combined data, or at a short-ISI ( $p = 0.617$ ) and medium-ISI ( $p = 0.059$ ). However, a significantly greater priming effect is found for the Ph condition compared to M at a Long-ISI ( $p = 0.028$ ).

The results show that priming effects for pseudo-derived (M) pairs can be obtained in the auditory modality. Experiment 2 investigates whether the effect in the M condition is morphological or phonological in nature.

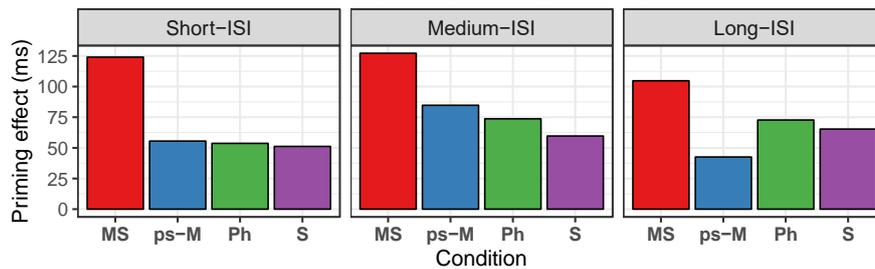


Figure 1. Priming effects (in ms) in Experiment 1. Priming effects reflect the difference between the RT to a target preceded by a related versus unrelated prime.

## Results Experiment 2

We analyzed log-transformed RTs with lme models. The results at Immediate Distance replicate the results at a medium-ISI in Experiment 1. The difference in priming between M and Ph was significant ( $p = 0.006$ ). With One Intervener, we find significant priming only for MS ( $p < 0.001$ ), but not for M ( $p = 0.312$ ), Ph ( $p = 0.051$ ), and S ( $p = 0.110$ ). No difference is found between the priming effects in MS and M ( $p = 0.078$ ), and M and Ph ( $p = 0.517$ ).

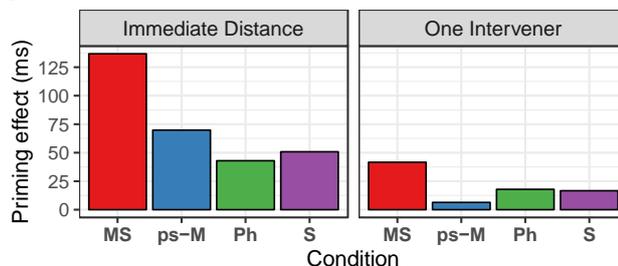


Figure 2. Priming effects (in ms) in Experiment 2.

## General discussion

The results show priming effects in all conditions, including the pseudo-M condition. This suggests that pseudo-derived words prime their ‘stem’ in the auditory modality, and that priming effects for pseudo-derived word pairs are therefore not merely an artefact of orthographic representations. However, the effect for pseudo-M is smaller than in the MS condition (in line with Feldman et al. 2015), and not consistently different from the phonological effect in Ph (in line with Milin et al. 2017). The lack of priming for pseudo-M with an intervener suggests that the effect is form-driven, and not morphological. In addition, our results have implications for theories of auditory word recognition. Specifically, the results in our Ph condition suggests that cohort competitors (e.g., *dog* to *dogma*, and *pill* to *pillow*) may stay activated longer than previously thought, which would explain the priming effects in the pseudo-derived condition as well.

## Acknowledgements

Many thanks are due to Meredith Tamminga, Robert J. Wilder, Amy Goodwin Davies, Hezekiah Akiva Bacovcin, and Kobey Shwayder.

## References

- Beyersmann, E., Ziegler, J.C., Castles, A., Coltheart, M., Kezilas, Y., Grainger, J. 2016. Morpho-orthographic segmentation without semantics. *Psychonomic Bulletin and Review*, 23(2), 533–539.
- Feldman, L.B., O’Connor, P.A., del Prado Martín, F.M. 2009. Early morphological processing is morphosemantic and not simply morpho-orthographic. *Psychonomic Bulletin and Review*, 16(4), 684–691
- Feldman, L.B., Milin, P., Cho, K.W., Moscoso del Prado Martín, F., O’Connor, P. A. 2015. Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in human neuroscience*, 9, 111.
- Longtin, C.-M., Segui, J., Halle, P. 2003. Morphological priming without morphological relationship. *Language and Cognitive Processes*, 18(3), 313–334.
- Marslen-Wilson, W. D., Bozic, M., Randall, B. 2008. Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes*, 23(3), 394–421.
- Milin, P., Feldman, L.B., Ramscar, M., Hendrix, P., Baayen, R.H. 2017. Discrimination in lexical decision. *PLoS one*, 12(2), e0171935.
- Rastle, K., Davis, M.H. 2008. Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, 23(7–8), 942–971.
- Rastle, K., Davis, M.H., New, B. 2004. The broth in my brother’s brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review*, 11(6), 1090–1098.

# Neurodegeneration and the Gothenburg MCI study

Marie Eckerström, Anders Wallin

Institute of Neuroscience and Physiology, Sahlgrenska Academy, Sweden

DOI:

## Abstract

Mild cognitive impairment (MCI) is the stage between the expected cognitive decline of a normal aging population and the more serious decline of dementia. Typically, the performance of a cognitive test is considered *impaired* if it is  $\Rightarrow$ 1.5 standard deviations below normal controls. The ongoing Gothenburg mild cognitive impairment (MCI) study is an attempt to conduct longitudinal in-depth phenotyping of patients with different forms and degrees of cognitive impairment using various assessment tools. This paper presents the background of the linguistic and extra-linguistic parameters for early detection of cognitive impairment-project, a research which is partly based on data acquired at the Gothenburg MCI-study.

Key words: neurodegeneration, mild cognitive impairment, dementia

## Introduction

Neurodegeneration is a progressive loss of the structure and function of neurons, eventually leading to neuron death. Some of the most common neurodegenerative disorders include Alzheimer's dementia (AD), frontotemporal dementia (FTD) and its variants, progressive supranuclear palsy (PSP), corticobasal degeneration (CBD), Parkinson's disease (PD), dementia with Lewy bodies (DLB), multiple system atrophy (MSA), and Huntington's disease (HD) (Erkkinen et al., 2018). Common factors of these diseases are that they gradually progress, and that there is to date no cure. The aetiology of the neuron death varies between diseases – it may be caused by tumours, vascular incidents such as stroke, viruses, toxins or excessive alcohol exposure, or the cause may be unknown. Partly as a result of a globally increased life expectancy, the prevalence of neurodegenerative disorders is increasing (Heemels, 2016).

Mild cognitive impairment (MCI) is a central concept in dementia, and especially Alzheimer's disease dementia research. MCI is not a disease in itself, but is a term used to describe an intermediate phase between normal cognition and dementia, characterized by cognitive impairment that is more pronounced than expected in normal aging but not at the advanced level of dementia (Petersen et al., 1997). Subjective cognitive impairment (SCI) is a related concept, although the cognitive decline in SCI is merely self-reported and not possible to measure with neuropsychological tests. The hypothesis

forming the basis for these concepts is that neurodegenerative dementia disorders start gradually, with a starting-point of neuropathological events decades before clear symptoms appear (preclinical phase), a first symptomatic phase of several years when the change is only detectable by the affected individual (SCI), a second symptomatic phase of several years when cognitive changes are still subtle but has become possible to detect using tests and also noticeable by others (MCI), followed by the dementia stage in which cognitive decline is at a level that reduces the possibility to live an independent life.

### **The Gothenburg MCI study**

The Gothenburg mild cognitive impairment (MCI) study started in 1999, and is a single-center clinical-pathophysiologic study aimed at investigating early and manifest phases of AD and vascular dementia (VaD) in patients seeking medical care for their cognitive decline (Wallin et al., 2016). The clinical setting for the study is an outpatient specialised memory clinic in the Gothenburg region in the south-west part of Sweden. The catchment area includes approximately 1 million residents. Patients are mainly referred to the clinic from primary care level, in cases when a dementia disorder is suspected but when the methods of basic examination in the primary care is not sufficient for diagnosis. As of 2019, 781 patients and 124 healthy controls were included in the study, and 233 converted to dementia.

One of the central objectives of the study is to contribute to a better foundation for future treatment options by enhancing the knowledge of disease processes that may cause cognitive impairment, using multimodal markers. One important step is to improve the characterization of the early and manifest phases of different dementia disorders. For this purpose, single and multimodal measurements from the clinical-, neuropsychological-, genetic-, biochemistry-, neurochemistry, and neuroimaging fields are employed (Wallin et al., 2016).

Neuropsychological tests are the gold standard to measure cognitive functioning, and is therefore central in dementia research. The neuropsychological test battery used in the Gothenburg MCI study includes 20 tests related to the cognitive domains of learning/memory; speed/attention, language function; visuospatial function and executive function.

### **Current status and the future of the Gothenburg MCI**

To date, several publications have been based on the neuropsychological tests in the Gothenburg MCI study. In the first years of the study, neuropsychological measures were analysed in 'broad strokes' using total scores and composite scores. The results generally highlight the importance

and validity of neuropsychological measures in dementia research. To summarize key publications, we reported that: SCI/MCI is a highly heterogeneous diagnostic entity with significant impairments in all cognitive domains (Nordlund et al., 2005); SCI/MCI subjects with vascular disease performed worse on tests measuring speed/attention, visuospatial function, and executive function (Nordlund et al., 2007); SCI/MCI with AD markers performed overall worse, especially on tests of speed/attention and episodic memory (Nordlund et al., 2008); a combination of vascular disease and AD biomarkers was associated with executive impairment (Nordlund et al., 2011); memory, visuospatial, and language symptoms characterized incipient AD (Nordlund et al., 2010); executive and speed/attention symptoms characterized incipient SVD (Nordlund et al., 2010); and neuropsychological tests were the best predictors of dementia with a combination of markers further improving the predictive ability (Eckerström et al., 2013).

The collaboration with Dr. Dimitrios Kokkinakis and colleagues within the project ‘Linguistic and extra-linguistic parameters for early detection of cognitive impairment’ (Kokkinakis et al., 2016) is an important next step to further refine the study of cognitive functioning in relation to dementia progression, using novel technical and analytical tools other than traditional neuropsychological testing; see for instance Fraser et al. (2017); Themistocleous et al. (2018) and Linz et al., (2019).

## References

- Eckerström, C., Olsson, E., Bjerke, M., Malmgren, H., Edman, Å., Wallin, A., Nordlund, A. 2013. A Combination of Neuropsychological, Neuroimaging, and Cerebrospinal Fluid Markers Predicts Conversion from Mild Cognitive Impairment to Dementia. *J of Alzheimer’s Disease* 36. 421–431. DOI 10.3233/JAD-122440.
- Erkkinen, M.G., Kim, M.O., Schwinn, M.D. 2018. *Clinical Neurology and Epidemiology of the Major Neurodegenerative Diseases*. Cold Spring Harbor Perspect Biol. 10 (4). pii: a033118. doi: 10.1101/cshperspect.a033118.
- Fraser, K.C., Lundholm Fors, K., Kokkinakis, D., Nordlund, A. 2017. An analysis of eye-movements during reading for the detection of mild cognitive impairment. Conference on Empirical Methods in Natural Language Processing (EMNLP). Pp. 1016–1026. Copenhagen, Denmark
- Heemels, M-T. 2016. Neurodegenerative disease. *Nature*. 10;539(7628):179. doi: 10.1038/539179a.
- Kokkinakis, D., Lundholm Fors, K., Nordlund, A. 2016. Data Resource Acquisition from People at Various Stages of Cognitive Decline - Design and Exploration Considerations. 7th International Workshop on Health Text Mining and Information Analysis (LOUHI). An EMNLP Workshop. Texas, USA.
- Linz, N., Lundholm Fors, K., Lindsay, H., Eckerström, M., Alexandersson, J., Kokkinakis, D. 2019. Temporal Analysis of Semantic Verbal Fluency Tasks in

- Persons with Subjective and Mild Cognitive Impairment. Sixth Workshop on Computational Linguistics and Clinical Psychology: Reconciling Outcomes. Minneapolis, USA.
- Nordlund, A., Rolstad, S., Hellström, P., Sjögren, M., Hansen, S., Ilin, A. 2005. The Goteborg MCI study: mild cognitive impairment is a heterogeneous condition. *J Neurol Neurosurg Psychiatry*. 76(11):1485-90.
- Nordlund, A., Rolstad, S., Klang, O., Lind, K., Hansen, S., Ilin, A. 2007. Cognitive profiles of mild cognitive impairment with and without vascular disease. *Neuropsychology*. 21(6):706-12. PMID: 17983284.
- Nordlund, A., Rolstad, S., Klang, O., Lind, K., Pedersen, M., Blennow, K., Edman, A., Hansen, S., Ilin A. 2008. Episodic memory and speed/attention deficits are associated with Alzheimer-typical CSF abnormalities in MCI. *J Int Neuropsychol Soc*. 14(4):582-90. doi: 10.1017/S135561770808079X.
- Nordlund, A., Göthlin, M., Ilin, A. 2011. Vascular disease, Alzheimer's disease biomarkers and cognition in mild cognitive impairment: additive or synergetic effects? *Dement Geriatr Cogn Disord*. 32(4):250-6. doi: 10.1159/000334653.
- Nordlund, A., Rolstad, S., Göthlin, M., Edman, A., Hansen, S., Ilin, A. 2010. Cognitive profiles of incipient dementia in the Goteborg MCI study. *Dement Geriatr Cogn Disord*. 30(5):403-10. doi: 10.1159/000321352.
- Petersen, R.C., Smith, G.E., Waring, S.C., Ivnik, R.J., Kokmen, E., Angelos, E.G. 1997. Aging, memory, and mild cognitive impairment. *Int Psychogeriatr*. 9 Suppl 1:65-9.
- Themistocleous, C., Eckerström, M., Kiknakis, D. 2018. Diagnosis of Mild Cognitive Impairment from Speech in Swedish using Deep Sequential Neural Networks. *Frontiers in Neurology*, section Neurodegeneration. doi: 10.3389/fneur.2018.00975.
- Wallin, A., Nordlund, A., Jonsson, M., Lind, K., Edman, Å., Göthlin, M., Stålhammar, J., Eckerström, M., Kern, S., Börjesson-Hanson, A., Carlsson, M., Olsson, E., Zetterberg, H., Blennow, K., Svensson, J., Öhrfelt, A., Bjerke, M., Rolstad, S., Eckerström, C. 2016. The Gothenburg MCI study: Design and distribution of Alzheimer's disease and subcortical vascular disease diagnoses from baseline to 6-year follow-up. *J Cereb Blood Flow Metab*. 36(1):114-31.

# Corpora and language variation in Greek

Fatima Eloeva<sup>1</sup>, Maxim Kisilier<sup>2</sup>, Olga Nikolaenkova<sup>3</sup>

<sup>1</sup>Department of Classical Philology, Vilnius University, Lithuania

<sup>2</sup>Hellenic institute, Saint Petersburg State University, Russia

<sup>3</sup>Department of General Linguistics, Saint Petersburg State University, Russia

DOI:

## Abstract

It seems that there are certain linguistic situations ideally adapted for the usage-based approach and the Greek case is one of them. This is a corpus-based study, which is based on the analysis and processing of a large variety of Greek texts in everyday spoken interactions. As our starting point, we argue the panchronic character of Greek lexicon, its extreme conservatism and the man-made character of the formation of the Greek literary standard. Some practical issues such as the choice between monotonic and polytonic orthography, lexemes tagging to obtain more data are addressed here.

Key words: language corpus, diglossia, modern Greek, language variation.

## Introduction

It seems that there are certain linguistic situations ideally adapted for the usage-based approach and the Greek case is one of them. Anyhow up to the present moment, Modern Greek does not possess its complete **electronic corpora** like most European languages.

There is no doubt that the Greek language question should not be neglected while creating linguistic corpora of Modern Greek. At the moment there are at least three corpora of Modern Greek but none of them currently reflects the existence of Greek diglossia.

The Hellenic National Corpus (<http://hnc.ilsp.gr/>) was created in the Institute for Language and Speech Processing in Athens. It consists of more than 47 million words in various types of texts. However, the texts with any peculiarities (for example dialectisms) are not included as well as the texts written before 1990. It means that the Hellenic corpus does not take into account Modern Greek literature.

The Corpus of Spoken Greek is a part of the Greek talk-in-interaction and Conversation Analysis research project, directed by prof. Pavlidou in the Institute of Modern Greek Studies at the Aristotle University of Thessaloniki ([http://ins.web.auth.gr/index.php?option=com\\_content&view=article&id=626&Itemid=251&lang=en](http://ins.web.auth.gr/index.php?option=com_content&view=article&id=626&Itemid=251&lang=en)). It was originally designed for the qualitative analysis of language and linguistic communication, especially from the perspective of Conversation Analysis. The material has been drawn from

naturally-occurring circumstances of communication and comprises everyday conversations among friends and relatives, classroom conversations, telephone calls and various types of TV broadcasts.

## Material

In the present paper, we introduce Corpus of Modern Greek ([http://web-corpora.net/GreekCorpus/search/?interface\\_language=en](http://web-corpora.net/GreekCorpus/search/?interface_language=en)). It consists of approx. 35.7 million tokens — but it is not restricted to any specific type of texts. The majority of texts incorporated in Corpus of Modern Greek come from contemporary Greek newspapers (Η Καθημερινή, Μακεδονία, Το Βήμα, Ελευθεροτυπία). However, there are also fiction, poetic, official, scientific, and religious texts, both original and translated, that were created in the 19<sup>th</sup> or in the 20<sup>th</sup> centuries. Thus, both the authors often regarded as supporters of Katharevusa, like Papadiamandis or Viziinos and demoticists (Nikos Kazantzakis) are present (Arkhangelskiy & Kisilier, 2018).

All texts have been morphologically annotated. It means that each word is provided with a lemma (dictionary form) and a set of morphosyntactic tags (such as case values, number values, etc.) which can be used in a search query. Morphological annotation was carried out with the help of a digital grammatical dictionary and a morphological analyzer (UniParser). Unfortunately, corpus still has no disambiguation, i. e. each word was annotated with all possible morphological analyses in all contexts (cf. Kuzmenko & Mustakimova, 2015). Another peculiarity of this corpus is a dictionary module that enables translation from English into Greek.

It is impossible to avoid the Greek language issue in the corpus since a variety of literary and non-literary texts is involved. The easiest solution was to tag “Katharevusa” or “Demotic Greek” authors or texts. Thus, for example, Alexandros Papadiamandis and his oeuvres were tagged as “Katharevusa” while Nikos Kazantzakis should be tagged as “Demotic Greek”. Recent corpus-oriented studies in Katharevusa Greek (Yakovleva, 2017) clearly demonstrated that this approach is inapplicable to non-Demotic material. We are going to illustrate it using the example of the renown Greek writer Alexander Papadiamandis (1851–1911) whose language is pretty often described as Katharevusa. This statement is mostly based on multiple archaic features in morphology or word formation:

τήν διήγησιν ‘story’ (accusative) vs. την διήγηση

έσηκώθη ‘[he] rose’ vs. σηκώθηκε

ἦσαν ‘[they] were’ vs. ήταν

and vocabulary:

ὕδωρ ‘water’ vs. νερό

οὖς ‘snake’ vs. φίδι

ωρίζω

A thorough analysis of his texts shows that there are a lot of lexical borrowings from

Turkish — ταμάμ ‘okay’ (< Turk. *tamam*)

Albanian — τσόπα ‘girl’ (< Alb. *çupë*)

Slavic — βάλτος ‘swamp’ (< Slav. *blato*)

along with dialectal and vernacular features:

πλιά ‘more’ vs. πιά

γρουνίζω ‘[I] know’ vs. γνωρίζω

Evidently, it would be a mistake to treat these words and morphological forms as Katharevusa, and if Papadiamandis and his texts are tagged as “Katharevusa”, all forms and words included in these texts will be compulsory ascribed to Katharevusa and may lead the user to various false conclusions concerning the language. We are sure that corpus may not decide for the user but it should provide enough information to let the user make his own decision.

Corpus of Modern Greek is now being moved to the new platform able to provide whole morphological paradigm with different statistic data, it will have a possibility for a dictionary mode and simultaneous representation of textual and audio material, later there shall be a module for translating from Russian into Greek, etc.

## Conclusions

Working on a new approach to the Greek language issue implementation includes two rather different options:

First of all, it is important to have a choice between monotonic and polytonic orthography. This choice formally exists even today, but it is supported only by few texts in polytonic. Introduction of new polytonic texts

will certainly face difficulties in their recognition, so special software is required.

The second option has to do with a new way of tagging. We suggest that instead of texts or authors one should tag lexemes, paradigms or even separate forms. The opposition should not be binary (“Demotic” vs. “Katharevusa”) but a triple one: “Demotic” vs. “Archaic” and “General” for the units that do not differ in Katharevusa and Demotic Greek. This kind of tagging will give the user new advanced search tools like choice of flexions or inflexion types from the point of view of their stylistic properties.

As a result, the user will have quantitative data for the texts he intends to analyze and he will be able to decide himself to which language or stylistic variant this text should be attributed.

### **Acknowledgements**

Corpus of Modern Greek compiled by Timofey Arkhangelskiy and Maxim Kisilier within the program “Corpus Linguistics” of the Russian Academy of Sciences. New platform design has been performed by Timofey Arkhangelskiy with Aleksander Rusakov and Maxim Kisilier financial support.

### **References**

- Arkhangelskiy T.A., Kisilier M.L. 2018. Greek corpora: achievements, goals and prospects. (Корпуса греческого языка: достижения, цели и задачи) *Indoeuropean linguistics and classical philology (Индоевропейское языкознание и классическая филология)* XXII. 1, 50–59.
- Kuzmenko E., Mustakimova E. 2015. Automatic disambiguation in the corpora of Greek and Yiddish. Annual International Conference “Dialogue” (2015, 27–30.05.2015), Moscow
- Yakovleva A.V. 2017. Ablative and allative marking of spatial position in katharevousa: corpus study (Аблативное и аллативное маркирование положения в пространстве в кафаревусе: корпусное исследование). Thesis presented for Master degree, Language theory and computational linguistics department, National Research University Higher School of Economics.

# Stream analysis for detecting stuttering episodes

Fabio Fassetti<sup>1</sup>, Ilaria Fassetti<sup>2</sup>, Simona Nisticò<sup>1</sup>

<sup>1</sup>DIMES, University of Calabria, Italy

<sup>2</sup>Therapeia Rehabilitation Center, Italy

DOI:

## Abstract

Stuttering is a communication disorder where a person is not able to speak fluently. A fluency disorder causes problems with the flow, rhythm and speed of speech. If one stutters, its speech may sound interrupted or blocked, as though the individual is trying to say a sound but it doesn't come out. This work is designed to help specialists in the evaluation of stuttering and recognize occurrences of disfluency episodes such as repetitions of sounds, syllables or words, silent pauses, hesitations or blocks during the speech.

Key words: Stuttering, stream detection, machine learning

## Introduction

Stuttering is a disorder of speech motor control, but the disorder of stuttering is more than just the stuttering it also involves a lifetime of dealing with the anxiety and avoidances caused by the stuttering. Typical disfluencies are: a whole word or phrase repeated three time or less; interjections and/or revisions in speech. Conversely, less-typical disfluencies are: any word that is repeated as part of a word or an initial sound, whole words or phrases repeated four or more times, prolonging one sound of a word, “blocking”, and any associated tension or effort observed in speech production. A fluency disorder can be evaluated by a speech-language pathologist which will conduct a precise evaluation with a series of tests, observations, interviews and listening. This work aims at helping the evaluation of stuttering and several types of disfluency. The proposed approach is based on deep learning and stream analysis and its goal is to automatically detect the points where verbal output is influential, thus helping in the early classification of stuttering or cluttering problems, providing the number of disfluencies and time intervals in which the disfluencies occur.

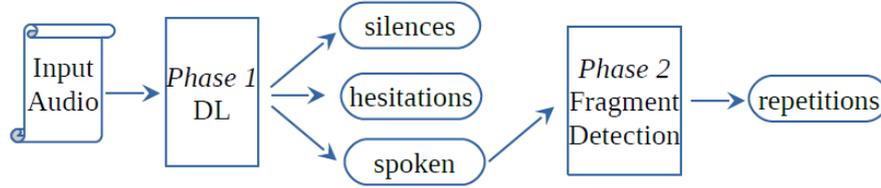
## Preliminaries

In the proposed work, an audio stream  $S$  in *wav* format and in mono mode is assumed as input, if the audio is in stereo mode the mean value between the two channels is considered. A window  $W$  in a stream  $S$  is the portion  $S(t_1, t_2)$  of  $S$  between instants  $t_1$  and  $t_2$ . The *energy* of a window  $W$  is the inner

F. Fassetti<sup>1</sup>, I. Fassetti, S. Nistico

product  $\langle W, W \rangle$ . The proposed framework makes use of a deep learner. Due to lack of space, details about deep learning theory are not provided.

Framework architecture:



The framework consists in two main phases. The first one aims at extracting relevant fragments from the input audio file. The goal is to remove intervals of silence and of hesitations. Hesitations are very difficult to distinguish from silence since they consist in spluttered letters/parts of words of very low energy.

The spoken fragments are then concatenated to obtain a clean stream provided as input of the second phase. During this phase, the goal is to individuate repetitions in form of very similar subsequent fragments in the clean stream. Next details about the detection technique are provided.

### Detection technique

Let  $F$  be a fragment of a stream  $S$ . Three features are extracted from  $F$ : ( $f_s$ ) portion of  $S$  associated with  $F$ , ( $f_f$ ) the spectrum of  $F$ , ( $f_c$ ) the MFCC of  $F$ . The similarity  $\Sigma(F_i, F_j)$  between two fragments  $F_i$  and  $F_j$  of a stream  $S$  is

$$\Sigma(F_i, F_j) = 1 - (w_s \cdot d_s(F_i, F_j) + w_f \cdot d_f(F_i, F_j) + w_c \cdot d_c(F_i, F_j)),$$

where  $w_s$ ,  $w_f$  and  $w_c$  are weights to be tuned and  $d_s$ ,  $d_f$  and  $d_c$  are three distance measures based on the three features as detailed in the following.

### Stream distance ( $d_s$ )

The distance  $d_s(F_i, F_j)$  between  $F_i$  and  $F_j$  according to feature  $f_s$  is obtained by a novel notion of distance based on the *Levenshtein* distance. This is due to the fact that, in order to make two windows comparable, the associated signal has to be aligned in a way such that the similarity is maximized. The *Levenshtein* distance is defined on strings and evaluates the number of edit operations to make the two strings equals. It uses three kinds of operations: *insert*, *delete* and *substitute* and each of them has a cost of 1. For example, the distance between “*sitting*” and “*kitten*” is 3 since the Stream analysis for detecting stuttering episodes substitution of  $s$  in  $k$ , of  $i$  in  $e$  and the removal of  $n$  have to be performed. In other words, the two strings are aligned for

minimizing the cost. Following the same approach, the aim is to align the fragments so that the Euclidean distance between them is minimized. Thus, in order to compute the distance between  $F_i$  and  $F_j$ , the insertion operation of element  $h$  costs  $F_i(h)^2$  the deletion operation of element  $h$  costs  $F_j(h)^2$  the substitution operation of element  $h_x$  of  $F_i$  and the element  $h_y$  of  $F_j$  costs  $(F_i(h_x) - F_j(h_y))^2$ .

### Spectrum-based distance ( $d_F$ )

The distance  $d_F(F_i, F_j)$  between  $F_i$  and  $F_j$  according to feature  $f_F$  is obtained by computing the *Fast Fourier Transform* (FFT) of the signal associated with  $F_i$  and  $F_j$ , by normalizing the power and, then, by computing the Euclidean distance between the two spectra.

### MFCC-based distance ( $d_C$ )

The distance  $d_C(F_i, F_j)$  between  $F_i$  and  $F_j$  according to feature  $f_C$  is obtained by computing the vectors composed by the *Mel-frequency cepstral coefficients* (MFCCs) and, then, by computing the Euclidean distance between these two vectors.

### Detection algorithm

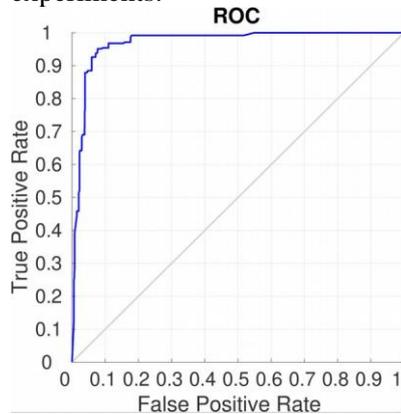
The *first phase* consists in recognizing and classifying spoken, noise, silence and hesitation fragments exploiting a multi-class deep-learner. The input stream is split in fragments of length  $\lambda = 0.5$  sec which are overlapped of  $\varepsilon = 0.25$  sec. With each fragment, a class is associated by the classifier and *hesitation fragments* are returned as *output*, *noisy* and *silence fragments* are removed and the *spoken fragments* compose the input of the second phase. Let  $S$  be the stream returned by the first phase, the goal of the *second phase* is to find windows  $W = S(t_w, t_w + ws)$  starting at time  $t_w$  of size  $ws$ , such that

1.  $\Sigma(W, S(t'_w, t'_w + ws)) < \alpha$  with  $t'_w > t_w$  and  $\alpha$  a threshold, and
2. Energy in  $S(t_w, t'_w)$  is negligible with respect to  $S(t_w, t_w + ws)$ .

Conditions are due to the fact that a repetition is characterized by a fragment  $F$  very similar to a subsequent fragment  $F_j$  (condition 1) interleaved by a fragment absent, namely  $t'_w = t_w + ws$ , or with spluttered letters, namely low energy signal (condition 2) which have not be filtered by phase 1.

## Experimental results

Performed experiments show that the approach is significantly accurated. The following figure reports the ROC curve associated with the output of the experiments.



The dataset employed for this preliminary campaign is publicly available [Howell et al. 2004] and it is composed by 152 audio wav files of stuttering English-speaking people of ages from 5 years to 47 years. Recordings have a length ranging from  $\approx 65$  s to  $\approx 1028$  s with a mean length of  $\approx 166$  s. Domain experts prepared the datasets by listening and manually selecting intervals where silences, hesitations and stuttering episodes occur.

## References

- Davis, S.B., Mermelstein, P. 1990. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. In: *Readings in Speech Recognition*, pp. 65–74. Morgan Kaufmann Publishers Inc.
- Howell, P., Davis, S., Bartrip, J., Wormald, L. 2004. Effectiveness of frequency shifted feedback at reducing disfluency for linguistically easy, and difficult, sections of speech. *Stammering Res.* 1(3), 309–315.
- Ingemann, F., Mermelstein, P. 1975. Speech recognition through spectrogram matching. *J. Acoust. Soc. Am.* 57, 253–255.
- Iverach, L., et al. 2018. Comparison of adults who stutter with and without social anxiety disorder. *J. Fluency Disord.* 56, 55–68.
- Koedoot, C., Bouwmans, C., Franken, M.C., Stolk, E. 2011. Quality of life in adults who stutter. *J. Commun. Disord.* 44(4), 429–443.
- Markett, S., et al. 2016. Impaired motor inhibition in adults who stutter - evidence from speech-free stop-signal reaction time tasks. *Neuropsychologia* 91, 444–450.
- Sammut, C., Webb, G.I. (eds.) 2010. *Encyclopedia of Machine Learning*. Springer.
- Weir, E., Bianchet, S. 2004. Developmental dysfluency: early intervention is key. *CMAJ Can. Med. Assoc. J.* 170, 1790–1791.
- Yairi, E., Ambrose, N. 2013. Epidemiology of stuttering: 21st century advances. *J. Fluen. Disord.* 38(2), 66–87.
- Yairi, E., Ambrose, N.G., Paden, E.P., Throneburg, R.N. 1996. Predictive factors of persistence and recovery: pathways of childhood stuttering. *J. Commun. Disord.* 29(1), 51–77

# Vocal and facial expressions and meaning effects in speech expressivity

Mario Fontes, Sandra Madureira  
Pontifical Catholic University of São Paulo, Brazil  
DOI:

## Abstract

The objective of this work is to investigate the congruence between non-verbal and verbal cues in persuasive speech. The selected corpus comprises video excerpts in which artists from divergent political perspectives provide support to the minister of the Brazilian Supreme Federal Court. The research methodology comprises: annotation of the video excerpts; text analysis; automatic analysis of the speakers' facial expressions and emotions by means of the FaceReader; analysis of the vocal quality and prosodic settings by means of the VPA; acoustic analysis of the data by means of the ExpressionEvaluator (Barbosa, 2009); and multivariate statistical analysis, applying MFA in R, package FactorMinerR, to correlate qualitative and quantitative variables. Results indicate the interplay among facial and vocal prosodies and intended persuasiveness.

Key words: verbal and non-verbal language, gestural prosody, multimodality, persuasion, speech expressivity

## Introduction

The objective of this work is to investigate the congruence between non-verbal and verbal cues in persuasive speech. The non-verbal aspects are key factors in oral communication reinforcing the semantic content, contradicting the semantic content or adding extra information.

## Method

### Corpus

The selected corpus is composed of video excerpts in which famous artists from divergent political perspectives provide support to the minister of the Supreme Federal Court, using a same utterance "Resist, Minister", with two different meanings: "Resist to changing the decision about the conviction at the second instance" and "Resist to the pressure of the ones who want you to change the decision about the conviction at the second instance".

The choice of the corpus was motivated by three factors: the semantic and pragmatic content of the text; the quality of the videos; and the possibility of comparing the oral interpretation of two groups of artists having different opinions about a same subject.

## Subjects

The subjects are four male and four female artists. According to their political beliefs, they can be divided into two groups (two males and two females in each one). Some of the artists are pro and others are contra conviction at the second instance, that is, in favor or not of incarcerating people who, found guilty and again on appeal to higher courts, receive a five-year jail sentence at the second instance.

## The analysis of facial and vocal expressions

The analysis of the facial action unities was performed automatically with the use of the FaceReader and the vocal quality and prosodic settings by means of VPA (Laver and Mackenzie Beck, 2007). FACS and VPA systems are contrasted in Madureira and Fontes (2019).

The acoustic measures were automatically extracted by the ExpressionEvaluator script developed by Barbosa (2009) for Praat. The script extracts 12 acoustic measures related to fundamental frequency (f0), intensity, spectral tilt and Long Term Average Spectrum (LTAS).

## Statistical procedures

In order to correlate the quantitative and qualitative variables, the statistical factorial method called Multiple Factor Analysis (MFA) has been applied with FactorMinerR (Husson et al, 2013).

The MFA was used to study similarities among stimuli relative to the 34 research variables structured in the three groups studied in this work: Gc1 (FACS), Gc2 (ExpressionEvaluator measures) and Gq2 (VPA). All measures were normalized by z-score. In order to verify the correlation among the groups of variables the Pearson Lg and RV coefficients were used.

## Results

The application of the MFA allows the analysis of the variables in several dimensions. The variables analyzed were displayed in five dimensions. Table 1 exhibits the contributions of the variables in these dimensions. In Dimension 1 Gc2 (ExpressionEvaluator) is more representative, in Dimension and 3 it is Gc1 (FACS) and in Dimension 4 and 5 is Gq2 (VPA).

Table 1. Significant variables in the five dimensions of analysis.

Groups	Dim. 1	Dim. 2	Dim 3.	Dim 4.	Dim 5.
Gc1	31.5841	46.9774	47.8636	15.4637	34.8319
Gc2	38.0663	10.7865	17.1027	34.1301	21.8097
Gq2	30.3496	42.2361	35.0337	50.4063	43.3584



In Table 2, the Lg Pearson coefficient is provided. The Gc1 (FACS) and the Gq2 (VPAS) are the most representative as indicated by the MFA coefficient. The Gq2 (VPAS) e Gc1 (FACS) groups are related ( $RV=0,81$ ) and Gq1 é the one which is closer to the other groups ( $RV=0,94$ ).

Table 2. LG Pearson coefficient values.

Groups	Gc1	Gc2	Gq2	MFA
Gc1	2.5342	1.2434	1.9847	2.2332
Gc2	1.2434	1.5271	1.4019	1.6170
Gq2	1.9847	1.4019	2.3686	2.2304
MFA	2.2332	1.6170	2.2304	2.3566

## Conclusion

The analysis of the video excerpts revealed that the non-verbal features, were influential in separating the two groups analysed. Findings provided pieces of evidence for meaning effects not possible to be interpreted without the use of multidimensional analysis. This was thought to be fundamental to disclose the paralinguistic meanings of the persuasive discourse.

## References

- Barbosa, P. A. 2009. Detecting Changes in Speech Expressiveness in participants of a radio program. *Proceedings of Interspeech*, Brighton, UK, 1, 2155-2158.
- Husson, F., Josse, J., Lê, S., Mazet, J. 2013. *FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R (version 1.25) [R package]*.
- Laver, J., Mackenzie-Beck J. 2007. *Vocal Profile Analysis Scheme*. Edinburgh: Queen Margareth University College, Speech Science Research Centre.
- Madureira, S., Fontes, M.A.S. 2019. The analysis of facial and speech expressivity: tools and methods. In J.M. Lahoz-Bengoechea, R. Pérez Ramón (Eds.), *Subsidia. Tools and resources for speech sciences* (pp. 19–26). Málaga: Universidad de Málaga.

# Reading and Mild Cognitive Impairment

Kristina Lundholm Fors<sup>1</sup>, Malin Antonsson<sup>1</sup>, Dimitrios Kokkinakis<sup>1,2</sup>, Kathleen C. Fraser<sup>3</sup>

<sup>1</sup>Department of Swedish, University of Gothenburg, Sweden

<sup>2</sup>Centre for Ageing and Health, University of Gothenburg, Sweden

<sup>3</sup>Digital Technologies Research Centre, National Research Council Canada, Canada

DOI:

## Abstract

In the present study, we investigated the discriminatory power of eye-tracking features in distinguishing between individuals with mild cognitive impairment (MCI) and healthy controls (HC). The eye movements of the study participants were recorded at two different time points, 18 months apart. Using a machine learning approach with leave-one-out cross-validation, we were able to discriminate between the groups with 73.6 AUC. However, somewhat surprisingly the classification was less successful using data from the second recording session, which might be attributed to the non-static nature of cognitive status. Still, the outcome suggests that eye-tracking measures can be exploited as useful markers of MCI.

Key words: mild cognitive impairment, dementia, eye-tracking, reading.

## Introduction and background

Dementia is a neurodegenerative disorder that is becoming more prevalent as the global population ages. In order to provide the best care and enable the alleviation of symptoms through lifestyle changes and access to medication, it is important to identify individuals at risk for dementia at an early stage. In this study, we used eye-tracking while reading as a means to identify persons with mild cognitive impairment (MCI). MCI is defined as a mild state of cognitive decline, which can affect several cognitive domains such as episodic memory, executive functioning and linguistic ability, and persons with MCI are at higher risk of developing dementia (Gauthier et al. 2006). As reading processes are dependent on several cognitive processes related to language and short-term memory, we believe that capturing the reading process with eye-tracking will allow us to differentiate between persons with MCI and healthy controls. Previous work has shown the effectiveness of using eye-tracking to detect signs of dementia due to Alzheimer's disease (Lueck, Mendez, Perryman 2000, Fernández et al. 2013, Biondi, Fernandez, Castro, Agamenonni 2017), but very little work has examined the potential utility of this technology for detecting milder phases of impairment (Fraser, Lundholm Fors, Kokkinakis, Nordlund 2017). In this study, we examine the performance of machine learning classifiers trained on eye movements

collected at two different time points, on the task of discriminating between participants with MCI and healthy controls.

## **Method**

### **Participants**

The participants in this study are a subcohort of the study population reported in Kokkinakis, Lundholm Fors, Björkner & Nordlund, (2017), which describes a broader project on language-related markers of early-stage cognitive impairment. In the current analysis, we include only those participants who successfully completed the eye-tracking component of the protocol at two different time points; specifically, 15 participants with MCI and 23 healthy controls (HC). Near the time of the first data collection session, each participant underwent a detailed neuropsychological assessment. The groups differed significantly on the MMSE (a test of cognitive functioning with a maximum of 30; MCI 28.6, HC 29.7), but not on age (mean age in years: MCI: 71.3, HC: 67.6) or education (mean education in years: MCI 14.0, HC 13.7). All participants were native speakers of Swedish. The study was approved by the local ethical committee review board (reference number: 206–16, 2016).

### **Data**

The task consisted of reading two short texts (one silently and one aloud) while eye movements were recorded with an EyeLink 1000 Desktop Mount (sampling rate:1000 Hz). The participants performed the same task twice, with 18 months in between. The eye-tracking data were analysed with a focus on several aspects of reading: fixations, saccades, reading duration etc. Each reading sample was represented by thirteen features, as proposed by Biondi et al. (2017) and further described by Fraser et al. (2017).

### **Statistical analysis and machine learning classification**

To test whether the eye-tracking data could be used to detect MCI, a machine learning experiment was performed. We compared three different classification algorithms: naïve Bayes (NB), support vector machine (SVM) with RBF kernel, and logistic regression (LR) with L2 regularization, all implemented in Python using the Scikit-learn library (Pedregosa et al., 2011). Each classifier was evaluated using leave-one-out cross-validation, with area under the ROC curve (AUC) as the evaluation metric. Features were standardized according to the training set in each fold, and default hyper-parameters were used. Additionally, we performed two-tailed t-tests with Bonferroni corrections for multiple comparisons, in order to investigate which features differed the most between the MCI group and the HC.

## Results

The classification results (Table 1) are consistently better for Session 1 (i.e. the first recording), with the best result of 0.736 compared to 0.658 for Session 2. Statistically, no difference in either session was significant at the corrected level ( $p < 0.004$ ). However, in Session 1, two features differ on the  $p < 0.05$  level: the MCI group has a larger number of “unique” fixations (fixations that occur once on words that were skipped in the first pass) ( $p=0.014$ ), and the standard deviation of the duration of the first pass fixations is greater in the MCI group ( $p=0.023$ ). In Session 2, one feature differs on the  $p < 0.05$  level: the persons with MCI have a greater mean amplitude of saccades than the controls ( $p=0.046$ ).

Table 1: Machine learning classification results (AUC).

	Logistic Regression	SVM	Naïve Bayes
Session 1	0.713	0.736	0.728
Session 2	0.644	0.524	0.658

## Discussion

The classification results from Session 1 and Session 2 in this study show that classification accuracy decreases from Session 1 to Session 2, even though we use data from the same participants. This is somewhat surprising, as we would expect differences between the MCI group and HC group to increase over time, as cognitive deterioration progresses. However, previous longitudinal studies report that up to 44% of persons fulfilling the criteria for MCI may return to normal within a year (Gauthier et al., 2006). Conversely, the prevalence of MCI increases with age, which would suggest that with time, the HC group is more at risk of developing MCI.

While no features were significant after correcting for multiple comparisons, the features that differ significantly on the  $p=0.05$  level can be interpreted in the context of MCI. The tendency in the MCI group to skip words in the first pass and come back to them in a later pass is consistent with them producing a greater number of “unique” fixations (fixations that occur once on words that were skipped in the first pass), and the greater mean amplitude of saccades also suggests that the MCI group make larger movements within the text instead of moving consistently from one word to the next. This disruption in the reading process can be attributed to the impairments in short term memory and executive function seen in MCI. The feature analysis also implies that there is a greater amount of variability in fixation durations in the MCI group, suggesting a more irregular pace through the text. These findings are consistent with our previous work

(Fraser et al., 2017), in which we concluded that the MCI and HC groups differed significantly in the linearity of the reading process.

In sum, this study highlights how eye-movement data from a longitudinal MCI-study can be used for classification purposes. Although the decrease in classification accuracy at first seemed surprising, this might be explained by variability in the cognitive status in both groups of participants, and underscores the need for updated neuropsychological assessments when evaluating persons with MCI longitudinally. The participants in our study are currently undergoing renewed neuropsychological assessments, and in future work, we plan to incorporate these results to evaluate if they explain the variation in the results, and if they can be used for better prediction modelling.

## Acknowledgements

This work has received support from The Swedish Foundation for Humanities and Social Sciences (grant agreement no: NHS 14-1761:1).

## References

- Biondi, J., Fernandez, G., Castro, S., Agamenonni, O. 2017. Eye-Movement behavior identification for AD diagnosis. arXiv preprint arXiv:1702.00837.
- Fernández, G., Mandolesi, P., Rotstein, N.P., Colombo, O., Agamenonni, O., Politi, L.E. 2013. Eye movement alterations during reading in patients with early Alzheimer disease. *Investig. Ophthalmol. Vis. Sci.* 54, 8345–8352.
- Fraser, K.C., Lundholm Fors, K., Kokkinakis, D., Nordlund A. 2017. An analysis of eye-movements during reading for the detection of mild cognitive impairment. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 1016–1026.
- Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R., Ritchie, K., Broich, K., Winblad, B. 2006. Mild cognitive impairment. *Lancet* 367, 1262–1270.
- Kokkinakis, D., Lundholm Fors, K., Björkner, E., Nordlund, A. 2017. Data collection from persons with mild forms of cognitive impairment and healthy controls - infrastructure for classification and prediction of dementia. *Proceedings of NoDaLiDa*. 131; LiU Electronic Press;172–82.
- Lueck, K.L., Mendez, M.F., Perryman, K.M. 2000. Eye movement abnormalities during reading in patients with Alzheimer disease. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Duchesnay, E. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 2825-2830.

# Italian vowel production by Arabic speakers

Manuela Frontera<sup>1</sup>, Emanuela Paone<sup>2</sup>

<sup>1</sup>Department of Cultures, Education and Society, University of Calabria, Italy

<sup>2</sup>Department of Humanities, University of Calabria, Italy

DOI:

## Abstract

This research wants to explore the production of Italian L2 vowels by native Arabic adult speakers living in Italy. It focuses, moreover, on the weight exerted by motivation on L2 speech. Accuracy in Italian L2 production is tested on two vowel oppositions, /i e/ and /o u/. First, with the purpose of judging the magnitude of spectral differences produced between /i e/ and /o u/, Euclidean distances between each opposition are calculated and compared, both for the Arabic speakers and an Italian control group. Then, spectral distances between Italian L1 and L2 B1-B0 and B3-B2 values are computed and examined in correlation with Arabic speakers' motivational indexes.

Key words: Unstressed vowels, Italian L2, Arabic L1, Motivation.

## Introduction

Native-like phonological competence in the L2 is well known to be the most challenging achievement for new speakers, whose perceptual-productive processes are led and deeply filtered by their former language experience (Flege 1987; Best & Tyler 2007). Moreover, matters of typological markedness and great “distances” between systems may slow down and even hinder the acquisition process, as well as the age factor (Eckmann 2008; Major 2014). However, extensive research studies have demonstrated the further fundamental weight of extra-linguistic/psychosocial factors, such as attitudes and motivation, on SLA (Moyer 2007), factors that will be explored through this study. Accuracy in Italian L2 production is here tested on two vowel oppositions, /i e/ and /o u/, in final unstressed context. These oppositions have been chosen because Arabic varieties' phonology lack of vowels /e/ and /o/ and, in addition, because all short vowels suffer from reduction processes when in open syllable and final position (Al-Ani 1970).

So, this research will test two main hypotheses: dealing with motivated subjects, it is argued that native Arabic speakers are able to accurately produce the aforementioned non-native vowel oppositions in Italian L2, but differences in motivational levels may trigger alterations in spectral distances from the native norm.

## Research methods

### Subjects

The informants for this research are a control group of 6 native Italians (mean age=37 years) and 6 L1 Arabic speakers, 4 males and 2 females (mean age=22,3 years) from Egypt and Tunisia, learners of Italian L2, living in Italy since 1,8 years, on average. Each Arabic speaker is associated to a 0-to-3 linguistic motivation index extracted, together with other sociolinguistic information, through an ad hoc questionnaire by a previous research (Frontera & Romito 2018).

### Speech material

The subjects read a list of 60 words (15 words per vowel) where /i e o u/ are in CV final unstressed position, pronounced in a carrier phrase, “Dico X rapidamente/con calma” (*I say X rapidly/with calm*), where X is one of the target words elicited in random order. Recordings take place in a silent room, using an Edirol UA-25 audio device, an Audio Technica AT4033 condenser microphone, with a sample frequency of 44.100 Hz, 16 bit, mono.

### Analyses

All the target vowels are labelled starting from the second positive peak up to the last clearly visible pulse. Vowel frequency values (F0, F1, F2, F3), in their mid-point, and durations in *ms* are extracted and manually corrected in Praat, then formant values are converted to Barks. Given this data, for each vowel, global frequencies, duration means, and standard deviations are observed and compared between the two groups. With the purpose of judging the magnitude of spectral differences that subjects produced between /i e/ and /o u/, Euclidean distances are calculated between each opposition, both for the Arabic speakers and the control group. Consequently, Italian L2 spectral distances are compared to the native Italian norm. Hence, new Euclidean distances between Italian L1 and L2 B1-B0 and B3-B2 values are computed, and examined in correlation with Arabic speakers' motivational indexes.

### Results

Due to preliminary ANOVAs which showed no significant differences in vowel highness and frontness/backness related to speakers' gender, and former Arabic dialectal varieties (Egyptian vs. Tunisian), data from male and female speakers were combined, as well as those obtained by all Italian L2 speakers.

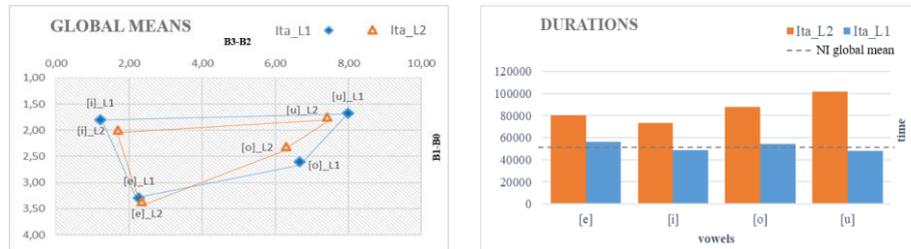


Figure 1 (on the left), Figure 2 (on the right). Global spectral means and durations of Italian /i e o u/ vowels produced by 6 Arabic speakers and a native Italian control group.

The ANOVA on Italian L1/L2 B1-B0 and B3-B2 values for the four vowels reveals significant differences between the two groups. More in depth comparisons (*t*-tests) show divergences in /i/ highness ( $p < .01$ ) and frontness ( $p < .001$ ), while /o/ and /u/ only seem to be, respectively, significantly higher ( $p < .001$ ) and slightly lower ( $p < .05$ ) when pronounced by the Arabic group (see Figure 1). Italian L1 vowel durations (in Figure 2) are systematically shorter in a significant way.

Repeated measures analyses carried out on the Euclidean distances obtained between /i/-/e/ and /u/-/o/ (see Table 1) confirm a significant variability between and within subjects, as regards the oppositions examined: a series of *t*-tests specifies that vowels /i e/ are significantly less distant in Italian L2 respect to the control group [ $t(185)=1,97, p < .01$ ].

Table 1. Mean spectral distances between /i/-/e/ and /u/-/o/ as produced by the Arabic speakers (Italian L2) and the control group (Italian L1).

	/i/-/e/	/u/-/o/
Italian L1	2,02 (0,36)	1,65 (0,73)
Italian L2	1,74 (0,45)	1,60 (0,82)

According to the Euclidean distances (ED) calculated between Italian L2 and Italian L1 /i e o u/, back vowels seem to be the most distant from the native norm indeed [/u/ M=1,56 (S.D.=0,98); /o/ M=1,33 (S.D.=0,75)] and their distance values diverge in a significant way ( $p < .001$  and  $p < .05$ , respectively) from those computed for front vowels /i/ and /e/ [/i/ M=1,01 (S.D. 0,6); /e/ M=1,04 (S.D. 0,6)]. Then, an ANCOVA is carried out to measure the impact exerted by motivation on vowels' EDs from the native models, using the Age of Arrival in Italy of each speaker, Lengths Of Residence and self-assessed levels of competence in Italian L2 as covariates. Motivational indexes (MIs) are scalar values on a continuum from 0 to 3 where the min. and max. measures are 1,88 and 2,59. Albeit the lowest index

corresponds to the highest mean ED ( $M=1,75$ , s.d. 1,23), the lowest ED ( $M=0,98$ , s.d. 0,6) is related to the MI 1,98, far from the max. ( $M=1,26$ , s.d. 0,6). If vowels confirm their strongest impact on variation, MIs trigger a slight significant variation of ED values (within subjects) ( $p<.01$ , eta squared 0,03): in order to observe the “direction” and strength of this effect, L2-to-L1 EDs, for each vowel, are finally correlated to Arabic speakers’ MIs, by computing Pearson’s correlation coefficients. Despite the resulting negative correlation between MI and distance scores, for /i/, /o/ and /u/, the obtained data is not statistically significant (p values are always  $>.05$ ).

## Conclusions

According to the present results, the Arabic speakers have shown themselves able to differentiate the tested oppositions: their spectral values are somehow centralized respect to the L1 model, but vowels seem to be properly discriminated by the investigated group. Though, front vowels are produced as consistently less distant from each other and, generally speaking, all vowel duration times almost double those of NIs: this could be interpreted as a sort of strategy of target achievement, to be better explored through further specific analyses, based on speech rate divergences. The motivational factor has not revealed a powerful correlation with better performances, nevertheless, in order to strengthen the present results, future analyses will be addressed to a broader number of non-native speakers and a wider range of MIs, including data from non or less motivated subjects.

## References

- Al-Ani, S. 1970. *Arabic Phonology*. Paris: The Hague.
- Best, C. T. & Tyler, M. D. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In Bohn, O.-S. and Munro M. (eds.) 2007. *Language Experience in Second Language Speech Learning*. In honor of Jim Flege, 13-34. Amsterdam: John Benjamins.
- Eckman, F. R. 2008. Typological markedness and second language phonology. In Hansen Edwards, J. G. and Zampini, M. L. (eds.) 2008. *Phonology and Second Language Acquisition*, 95-115. Amsterdam-Philadelphia: John Benjamins.
- Flege, J. 1987. The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15, 47-65.
- Frontera, M. & Romito, L. 2018. "Vorrei perdere il mio accento straniero": indici motivazionali di migranti culturali in Italia, *Quaderni del Dipartimento di Linguistica*, 26 (in press).
- Major, R. C. 2014. *Foreign accent: The ontogeny and phylogeny of second language phonology*. New York and London: Routledge
- Moyer, A. 2007. Do Language Attitudes Determine Accent? A Study of Bilinguals in USA. *Journal of Multilingual and Multicultural Development* 28, 502-518.

# Garden-path sentences and executive functions in normal aging

Larissa Rangel Ferrari<sup>1</sup>, Erica dos Santos Rodrigues<sup>1</sup>, Daniel Correa Mograbi<sup>2</sup>

<sup>1</sup>Department of Letters, Pontifical Catholic University of Rio de Janeiro, Brazil

<sup>2</sup>Department of Psychology, Pontifical Catholic University of Rio de Janeiro, Brazil

DOI:

## Abstract

This paper investigates how the elderly process temporarily ambiguous sentences and how syntactic processing difficulties can be related to the decline of executive functions, such as working memory and inhibitory control. The study consisted of the application of a psycholinguistic experiment and neuropsychological evaluation tests. The participants were 20 older adults and 20 younger adults, who had to resolve the ambiguity in garden-path sentences, such as “While the man hunted the deer ran into the woods” (Christianson et al. 2006). The results are congruent with the hypothesis that a decreased inhibition can make it more difficult for the elderly to implement syntactic reanalysis in conditions that favour good-enough processing.

Key words: comprehension, syntax, aging, executive functions, garden-path effect

## Introduction

Studies on language in normal aging have reported that the elderly tend to have more difficulty in understanding sentences of greater syntactic complexity, such as passives and sentences with embedded clauses or syntactic ambiguity (Kemper 1986; Christianson, Williams, Zacks, Ferreira 2006; Obler, Fein, Nicholas, Albert 1991; Zhu, Hou, Yang 2018). The literature has also shown a possible correlation between difficulties in language comprehension and the decline of non-linguistic cognitive functions, such as processing speed, working memory capacity, and the ability to sustain attention and inhibit irrelevant information (Hasher, Zacks 1988; Obler et al. 1991; Salthouse 1994, Kemper, Herman 2006).

In our study, we investigated how older adults process sentences with temporary syntactic ambiguity, which induce a garden-path effect, like “While the man hunted the deer ran into the woods” (Christianson et al. 2006). In this example, the noun phrase (NP) “the deer” is the point of ambiguity, because at first it can be interpreted as the object of the subordinate clause verb (“hunted”) or the subject of the main clause verb (“ran”) but when the end of the sentence is reached, the only grammatically acceptable reading is to interpret “the deer” as the subject. Thus, a reanalysis process has to be performed, suppressing the syntactic representation of the

critical NP as the object. The resolution of the temporary ambiguity seems to place greater demands on working memory and inhibitory control.

Christianson et al. (2006) examined how older adults interpret sentences with garden-path effect and found the elderly were generally more likely to endorse the incorrect interpretation under some conditions. They concluded that the elderly may rely more on heuristic-like good-enough processing in order to compensate for a decline of working memory capacity. However, the difficulty in revising the initial incorrect interpretation may also be linked to the decline of inhibitory mechanisms, since the parser needs to inhibit the first, incorrect interpretation in order to conduct reanalysis. We expanded this investigation in our research: we analysed the correlation between the participants' performance in ambiguity resolution and their performance on working memory and inhibition tests.

### **Our study**

We applied a psycholinguistic experiment (an off-line questionnaire) and three neuropsychological tests: Digit Span, to evaluate working memory, Stroop, to evaluate inhibitory control, and Mini-Mental State Examination (MMSE), to discard participants with dementia.

### **Participants**

The study was conducted with 40 Brazilian Portuguese speakers: 20 older adults (11 women, 9 men; mean age = 72.2) and 20 younger adults (9 women, 11 men; mean age = 27.3). All young adults were graduate or undergraduate students. The level of education of the elderly ranged from elementary school to college level. Only the elderly performed the MMSE.

### **Experiment**

In the psycholinguistic experiment, two variables were taken as within-subjects: (i) the distance between the critical NP and the verb of the main clause; and (ii) the plausibility of keeping the critical NP as the object of the subordinate clause verb. Examples of sentences in each condition follow:

- (1) short plausible sentence: “Enquanto o rapaz bebia a cerveja era derramada no balcão” (“While the young man drank the beer was spilled on the counter”);
- (2) short implausible sentence: “Enquanto a jovem bebia o vinho era mantido lacrado na adega” (“While the young woman drank the wine was kept sealed in the cellar”);

- (3) long plausible sentence: “Enquanto a mulher despertava o bebê que era tranquilo e sereno era picado por um inseto” (“While the woman woke up the baby who was quiet and serene was bitten by an insect”);
- (4) long implausible sentence: “Enquanto o homem despertava a mulher que era jovem e tranquila era agredida no trabalho” (“While the man woke up the woman who was young and quiet was assaulted at work”).

### Materials and procedure

Stimuli consisted of 32 experimental sentences and 28 fillers, randomly presented. There were 8 trials for each of the 4 conditions. The fillers consisted of active and passive sentences and with perspective verbs.

The sentences were presented separately on a notebook screen. Participants could only read each sentence twice. They had to answer questions with “yes” or “no” after reading each sentence. The questions related to whether the participant had retained the analysis of the critical NP as the object (e.g. “Did the man hunt the deer?” for the sentence “While the man hunted the deer ran into the woods”).

### Results

The data was analysed using the Statistical Package SPSS 23, and a variance analysis (ANOVA) was applied to the hit rate results. T-tests were used to compare the performance on the Digit Span test and Stroop test. The Pearson correlation coefficient ( $r$ ) was applied to identify possible correlations between the linguistic factors and the executive functions.

The elderly performed worse than young adults on the Digit Span test, on the direct digit order phase ( $p = .011$ ) and on the reverse digit order phase ( $p = .023$ ). On the third phase of the Stroop test, the elderly also presented worse results than the young adults ( $p = .011$ ). The task execution time for the Stroop test was significantly longer for the elderly ( $p < .001$ ).

The results showed a main effect of plausibility [ $F(1,38) = 87.37$ ;  $p < .001$ ] and a main effect of extension of the ambiguous region [ $F(1,38) = 18.61$ ;  $p < .001$ ]. An interaction effect was also observed between plausibility and group [ $F(1,38) = 7.99$ ;  $p = .007$ ]. From the pairwise comparisons, the difference in performance between the groups was only significant on the plausible condition ( $p = .035$ ). The elderly performed better on the implausible than the plausible conditions. In view of the differences in the participants' level of education, a covariance analysis (ANCOVA) was performed, with the inclusion of education as a covariate. The interaction effect between plausibility and group remained [ $F(1,37) = 9.46$ ;  $p = .004$ ].

A positive correlation was observed between age and number of errors in the third phase of the Stroop test ( $r = .35$ ;  $p = .05$ ); between age and task execution time in the Stroop test ( $r = .58$ ;  $p = .01$ ); and between age and task execution time in the third phase of the Stroop test ( $r = .68$ ;  $p = .01$ ). There was also a negative correlation between plausibility and number of errors in the third phase of the Stroop test ( $r = -.33$ ;  $p = .05$ ); between plausibility and task execution time in the third phase of the Stroop test ( $r = -.36$ ;  $p = .05$ ); between plausibility and age ( $r = -.44$ ;  $p = .01$ ); between age and performance in the direct digit order phase ( $r = -.40$ ;  $p = .05$ ); and between age and performance in the reverse digit order phase ( $r = -.38$ ;  $p = .05$ ) of the Digit Span test.

## Discussion

As we predicted, solving the ambiguity in the plausible and the long conditions was more complex for both groups of participants. However, the older participants performed significantly worse on the plausible condition and on the Stroop test than the younger participants.

These data and the correlation between plausibility and the performance on the Stroop test suggest that a decline in inhibition in aging may have hindered implementation of a full syntactic reanalysis in the plausible condition. The elderly may tend to implement a good-enough processing strategy, based on pragmatic knowledge. Moreover, the correlation between age and performance on the Digit test converges with the literature, which reports that aging is accompanied by a decline in working memory.

## References

- Christianson, K., Williams, C.C., Zacks, R.T., Ferreira, F. 2006. Younger and older adults' "good-enough" interpretations of garden-path sentences. *Discourse processes*, 42, 205-238.
- Hasher, L., Zacks, R. 1988. Working memory, comprehension and aging: a review and a new view. In Bower, G. K. (ed.). *The psychology of learning and motivation*. New York: Academic Press, 192-225.
- Kemper, S. 1986. Limitation of complex syntactic constructions by elderly adults. *Applied Psycholinguistics*, 7, 277-288.
- Kemper, S., Herman, R.E. 2006. Age differences in memory-load interference effects in syntactic processing. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* 61, 327-332.
- Obler, L.K., Fein, D., Nicholas, M., Albert, M.L. 1991. Auditory comprehension and aging: decline in syntactic processing. *Applied Psycholinguistics*, 12, 433-452.
- Salthouse, T.A. 1994. The aging of working memory. *Neuropsychology* 8, 535-543.
- Zhu, Z., Hou, X., Yang, Y. 2018. Reduced Syntactic Processing Efficiency in Older Adults During Sentence Comprehension. *Frontiers in Psychology* 9, article 243.

# Iconicity in grammatical variation

Takashi Fujiwara<sup>1</sup>, Fuminori Nakamura<sup>2</sup>, Daisuke Suzuki<sup>1</sup>

<sup>1</sup>Faculty of Foreign Studies, Setsunan University, Japan

<sup>2</sup>College of Liberal Arts and Sciences, Kitasato University, Japan

DOI:

## Abstract

This study examines the form-function relation of prepositional phrases, taking up the use of *of interest* as a sample case. It is well known that prepositional phrases such as this function like adjectives. By conducting a questionnaire-based survey, we investigated two predictions: (i) whether a formal or real subject filled the Subject slot in a sentence, and (ii) whether the expressions co-occurred with modal verbs. The results of the analysis demonstrate that both variables and their combination have evident effects on the choice between the two alternatives, and these can be interpreted in terms of the iconic principle. The results of this analysis contribute to ongoing discussions about iconicity theory and its applications to related fields.

Key words: prepositional phrases, synonymy, questionnaire-based survey, iconicity

## Introduction

This paper examines prepositional *of*-phrases and their synonymous expressions, paying particular attention to pairs of *of*-phrases and adjectives. According to Quirk et al. (1985: 732), prepositional phrases are nearly equivalent in meaning to adjective or noun phrases that function as complements in some cases, as shown in examples (1a, b):

- (1) a. That is of no importance.
- b. That is *unimportant*. (Quirk et al. 1985: 732)

Biber Johansson, Leech, Conrad (1999) and Huddleston and Pullum (2002) also make similar claims, indicating that *of*-phrases are grammatically equivalent to adjectivals; on the other hand, the literature does not provide a fine-grained description of the differences between the two. Despite the linguistic fact of the difference in form (i.e., grammatical category), the clear-cut functional differences between them remain unclear. Bolinger (1977: x) maintains that “the natural condition of a language is to preserve one form for one meaning, and one meaning for one form.” This study, therefore, extends our analysis that sheds light on the characteristics of *of*-phrases and uncovers the combination of two kinds of factors in the use

of an experimental method. In this analysis, we focused special attention on the synonymous pair of *of interest* and *interesting*.

## Methodology

For this study, a paper-based experiment was conducted using a 2×2 factorial design crossing the Subject Type factor and the Modality factor. These variables refer to (i) whether a formal or real subject filled the Subject slot in their clauses and (ii) whether or the expressions co-occurred with modal verbs (e.g., *may*, *must*, *will*), respectively. These two factors will prove significant in the following analysis for the distinction between *of*-phrases and their equivalents.

Multiple choice experimental designs require four options rather than two. For this reason, we added *interest* and *with interest*, both of which have something in common with the target expressions from a semantic or structural viewpoint, yielding the alternatives (1: *interest*, 2: *interesting*, 3: *of interest*, and 4: *with interest*). We conducted a cloze test containing 16 target fragments in the small text passages, adapted from the British National Corpus, as shown in Examples (2a–d):

- (2) a. [Formal / Non-Modal]  
It is ( ) to compare Pausanias's account of Polygnotos's Troy at Delphi with two earlier vase pictures of the Sack.
- b. [Formal / Modal]  
It **will** be ( ) to see whether these objectives are attained. As regards the first, it seems unlikely.
- c. [Real / Non-Modal]  
The 1983 and 1987 general elections were ( ) for psephological as well as for political reasons.
- d. [Real / Modal]  
A text of the past **may** be ( ) to a critic, a philosopher, and a historian, and each would discover his appropriate interest within it.

In this multiple-choice task, a questionnaire with 64 items (16 target fragments and 48 fillers) was administered to 30 subjects, 15 of whom were native speakers of American English and the other 15 were native speakers of British English (17 females, 13 males).<sup>1</sup> The subjects, aged 20–50 years, were asked to complete each sentence by choosing the most natural of the four expressions in the provided contexts as spontaneously as possible. As a result, we obtained a total of 480 sentence completion responses (16 fragments × 30 participants) and conducted further quantitative analysis on these responses in terms of frequency.

## Results and discussion

Table 1 exhibits the frequency of analyzable completions produced in this experiment. It is shown that *of interest* and *interesting* display the highest frequencies.

Table 1. Frequency and percentage of the four expressions in this experiment.

Expressions	Frequency	%
interest	2	0.4
interesting	278	57.9
of interest	196	40.8
with interest	4	0.8
Total	480	100.0

We also explored the interaction between the factors in more detail. Table 2 shows the percentage of analyzable occurrences when presented with the Formal or Real subject and with and without modal verbs. (Table 2 focuses on the frequencies of *interesting* and *of interest* with respect to each factor.) The important point displayed in Figure 1 is that the use of the two expressions fluctuates according to the four experimental conditions. A decreased use of *of interest* can be observed under the Formal subject conditions (the total of the Formal/Non-Modal condition and the Formal/Modal one). However, the preference for the phrase increases under the Real/Non-Modal condition; its occurrence becomes very close to *interesting*. When the Real subject co-occurs with a Modal, the preference for *of interest* is strongly established.

Table 2. Frequency of *interesting* and *of interest* in this experiment.

	interesting	of interest
Formal / Non-Modal	95	23
Formal / Modal	97	20
Real / Non-Modal	67	52
Real / Modal	19	101

Based on the results obtained from this experiment, two important implications follow. First, although the effect of *subject type* is greater than that of *modality*, each factor has a decisive effect on the choice between the paired items. Second, more importantly, these two factors reinforce one another in the use of synonymous expressions.

Another point that is evident from this study is that the participants' strong preference for *of interest* can also be seen in structures such as below:

- (3) a. What happens in childhood is ( ) ...  
b. The relative emphasis on neuroses in female GP referrals is ( )

...

The use of *of interest*, thus, indicates a close relationship to the specification of the subject and the expression of modality, and more information is encoded in the *of interest* clause than in that of *interesting*. In terms of form and meaning, it is plausible that these characteristics of the phrase *of interest* are derived from its form—that is, from the fact that it is a (prepositional) phrase, not a word (adjective). Simply put, such a complex use of *of interest* has a close affinity to its phrase structure. This situation corresponds precisely to the predictions of the hypothesis of iconicity of complexity (cf. Haiman 1980, 1985). This idea also adheres to Lakoff and Johnson's principle, "more of form is more of content" in our conceptual system (Lakoff and Johnson 1980: 127).

## Conclusion

To test the factors *subject type* and *modality* in isolation and in unison, an experimental technique was employed in this study. We generated four experimental conditions by manipulating these two factors. The results revealed a significant interaction between the factors in the choice of paired items, and we interpreted them from the iconic point of view.

## Notes

We also explored regional differences in the use of *of*-phrases and their equivalents; however, it did not reach standard levels of significance.

## References

- Biber, D., Johansson, S., Leech, G., Conrad, S. 1999. Longman Grammar of Spoken and Written English. Harlow, Pearson.
- Bolinger, D. 1977. Meaning and Form. London, Longman.
- Haiman, J. 1980. The iconicity of grammar: Isomorphism and motivation. *Language* 56, 515–540.
- Haiman, J. 1983. Iconic and economic motivation. *Language* 59, 781–819.
- Huddleston, R. and Pullum, G.K. 2002. The Cambridge Grammar of the English Language. Cambridge University Press.
- Lakoff, G., Johnson, M. 1980. *Metaphors We Live By*. University of Chicago Press.
- Quirk, R., Greenbaum, S., Leech, G., Svartvik, J. 1985. *A Comprehensive Grammar of the English Language*. London, Longman.

# Lexical deterioration in Alzheimer's disease

Frédérique Gayraud<sup>1</sup>, Isabelle Simoes Loureiro<sup>2</sup>, Camille Frouin<sup>1</sup>, Cynthia Collette<sup>2</sup>, Laurent Lefebvre<sup>2</sup>, Melissa Barkat-Defradas<sup>3</sup>

<sup>1</sup>Laboratoire Dynamique du Langage, University of Lyon & CNR, France<sup>2</sup>

<sup>2</sup>Cognitive Psychology and Neuropsychology dept, University of Mons, Belgium

<sup>3</sup>Institut des Sciences de l'Évolution de Montpellier, France

DOI:

## Abstract

This research aimed to assess the validity of the retrogenesis hypothesis applied to the lexico-semantic knowledge. According to this hypothesis, neurodegenerative mechanisms would reverse the order of acquisition in normal development. We present two studies comparing the development and loss of semantic knowledge. Study 1 compares the development of semantic knowledge about objects in children aged from 5 to 9 years old and its loss in Alzheimer's disease. Study 2 compares the performances of children and patients with Alzheimer's disease in a picture-naming task. Both studies exhibited mirrored evolution of semantic knowledge extending the retrogenesis theory to lexico-semantic knowledge.

Key words: retrogenesis, Alzheimer's disease, semantic knowledge, lexical access

## Introduction

Lexico-semantic network is the result of a progressive construction across lifespan. During childhood, the features of objects and their verbal labels are progressively stored to build our semantic knowledge and language. In AD, both sides, language and semantic knowledge, can be affected as demonstrated by word finding difficulties, circumlocutions, paraphasia and/or semantic errors.

The specific aim of this research is to analyze whether the understanding of lexico-semantic knowledge constitution in childhood brings clarification about its deterioration in AD. Retrogenesis theory (Reisberg, 1984) suggests that the earlier an ability is developed, the stronger it will resist to aging and to degenerative processes such as AD. Thus, one can expect a retrogenic trajectory of lexico-semantic aspects in AD manifested by the loss of semantic knowledge following a path opposite to their acquisition.

Some studies have highlighted results suggesting that this hypothesis of retrogenesis should be considered (e.g. Mello et al., 2008). The results of these studies show that the lexico-semantic aspects that are acquired later in development are also the most vulnerable in AD. The goal of the present study was then to test the retrogenic hypothesis by applying it to the lexico-semantic deterioration in normal aging and in AD.

## Study 1: Semantic knowledge

Our first study aimed to investigate the development of semantic knowledge about objects in children aged from 5 to 9 years old and to compare this acquisition to the semantic loss in Alzheimer's disease.

### Method

A total of 191 French native speakers were enrolled in the experiment. 93 children were allocated into 3 age groups: 5-year-olds ( $n=30$ , 17 girls); 7-year-olds ( $n=30$ , 20 girls) and 9-year-olds ( $n=33$ , 19 girls). 32 healthy elderly and 66 AD patients were also tested. AD patients were divided into three groups depending on their global cognitive deterioration measured by their performance at the Mini-Mental State Examination (MMSE): Mild (AD1:  $MMSE > 20$ ,  $n=20$ ); Moderate (AD2:  $16 < MMSE < 19$ ,  $n=16$ ), and Advanced stages (AD3:  $MMSE < 16$ ,  $n=30$ ). No difference of age and socio-cultural level appeared between the four elderly groups ( $p < .05$ ). The Semantic Knowledge Questionnaire (SKQ) was used for all participants. Initially proposed by Laiacona (1993), this tool was revised and adapted in French for children (Simoes Loureiro & Lefebvre, 2015a) and for AD patients (Simoes Loureiro & Lefebvre, 2015b). SKQ contains 120 questions about 30 objects with different levels of hierarchy and attributes in semantic memory (Q1: questions about general superordinate aspects, e.g. Is the ostrich an animal, a plant, or an object?; Q2: questions about intracategorical aspects, e.g. Is the ostrich a four-legged animal, a bird, or an insect?; Q3: questions about perceptual attributes, e.g. Is the ostrich's neck longer than a rooster's neck? ; Q4 : questions about thematical/functional attributes, e.g. Does the ostrich run, walk, or swim?).

### Results and discussion

ANOVA on repeated measures were carried out using the errors to each type of question (Q1, Q2, Q3, Q4) as 'within-subject factor,' and the groups as 'between-subject factor'.

Results indicated a group effect, ( $F=43.884$ ;  $p < .001$ ) demonstrating a decrease of the total number of errors to SKQ in childhood and an increase with the stages of AD, a type of question effect ( $F=230.083$ ;  $p < .001$ ) demonstrating fewer errors at the superordinate level (Q1+Q2) than at the subordinate level (Q3+Q4) and an interaction effect ( $F=23.066$ ;  $p < .001$ ) (for a more detailed description of the results, see Simoes Loureiro & Lefebvre, 2016).

Those results demonstrate an evolution/deterioration of the different levels of hierarchy of semantic knowledge with a slight improvement of general and intracategorical superordinate information from 5 to 9 years old

and a slight deterioration in AD, contrary to the subordinate level (perceptual and thematic/functional), with a later formation in childhood and an earlier deterioration in AD. These results sustain the hypothesis of Reisberg (1984). Indeed, we observed that for the knowledge tested in this study, the earlier the children acquired concepts, the later the AD patients lost it. Study 2 will consider the importance of the developmental trajectory of semantic knowledge by considering the age of word acquisition as a factor in anchoring this knowledge.

## **Study 2: Naming task**

The purpose of study 2 was to compare naming performances of children and patients with AD. Based on the assumption that early acquired words are more robust than late acquired ones, we predicted that the scores of patients at different stages of the disease and those of children at different developmental stages should show a mirror pattern, especially for late acquired words.

### **Method**

A total of 160 participants: 80 AD patients, 60 children and 20 elderly controls participated in the study. The AD group was split into four subgroups depending on the MMSE scores: mild (MMSE = 20-25), moderate (MMSE = 16-19), moderately severe (MMSE = 10-15), and severe (MMSE < 10). Similarly, the children were divided into 3 subgroups: early childhood (MMSE = 10-15, from 3 to 5 years old), middle childhood (MMSE = 16-19, from 5 to 7 years old) and late childhood (MMSE = 20-25, from 8 to 10 years old).

For the naming task, 24 pictures were selected from Rossion & Pourtois (2004): 12 pictures depicted words acquired early (mean AoA = 1,8 years), and 12 pictures corresponded to words acquired later (mean AoA = 2,8 years). Words in both lists were controlled for frequency, name agreement, word length, phonological neighborhood and visual complexity.

### **Results and discussion**

We found a significant effect of group ( $p < .0001$ ), AoA ( $p < .0001$ ), as well as a significant interaction ( $p = .001$ ). Patients with AD showed lower scores compared to other participants, early acquired words displayed higher scores than late acquired words, especially for children and AD patients. Focusing on early acquired words, Post hoc Tuckey analysis showed no significant differences between subgroups except for severe AD patients, who showed significantly lower scores ( $p < .0001$ ). Turning to late acquired words, the results showed a mirror pattern between the subgroups of AD patients and children.

Globally, our findings are in line with the retrogenesis hypothesis. We also found a strong AoA effect, which had however, a more detrimental effect in AD patients than in children.

## Conclusion

Our findings suggest that globally, the retrogenesis hypothesis is relevant and useful for a better understanding of language decline in AD, since developmental changes in the child are better known than changes occurring in AD. Nevertheless, it is also important to take into account the specificities of each population, as AD patients benefit from a cumulative life experience that should not be ignored (Moos, 2011).

## References

- Laiacona, M., Barbarotto, R., Capitani, E. 1993. Perceptual and associative knowledge in category specific impairment of semantic memory: A study of two cases. *Cortex*, 29(4), 727-740.
- Mello, C.B.D., Abrisqueta-Gomez, J., Xavier, G.F., Bueno, O.F.A. 2008. Involution of categorical thinking processes in Alzheimer's disease: Preliminary results. *Dementia & neuropsychologia*, 2(1), 57-62.
- Moos, I. 2011. Humour, irony and sarcasm in severe Alzheimer's dementia— a corrective to retrogenesis?. *Ageing & Society*, 31(2), 328-346.
- Reisberg, B., Ferris, S.H., Anand, R., de Leon, M.J., Schneck, M.K., Buttinger, C., Borenstein, J. 1984. Functional staging of dementia of the Alzheimer type. *Annals of the New York Academy of Sciences*, 435(1), 481-483.
- Reisberg, B., Franssen, E.H., Souren, L.E., Auer, S.R., Akram, I., Kenowsky, S. 2002. Evidence and mechanisms of retrogenesis in Alzheimer's and other dementias: management and treatment import. *American Journal of Alzheimer's Disease & Other Dementias®*, 17(4), 202-212.
- Rossion, B., Pourtois, G. 2004. Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33(2), 217-236.
- Simoes Loureiro, I., Lefebvre, L. 2015a. Développement d'un questionnaire de connaissances sémantiques des objets naturels et manufacturés pour enfants de 5 à 9 ans. *L'année Psychologique*, 115(3), 409-434.
- Simoes Loureiro, I., & Lefebvre, L. 2015b. Le QCS: Questionnaire de connaissances sémantiques pour déterminer le stade de détérioration sémantique chez les patients atteints de la maladie d'Alzheimer. *Gériatrie et Psychologie Neuropsychiatrie du Vieillissement*, 13(2), 225-233.
- Simoes Loureiro, I., Lefebvre, L. 2016. Retrogenesis of semantic knowledge: comparative approach of acquisition and deterioration of concepts in semantic memory. *Neuropsychology*, 30 (2) 1-7.

# Lexical variation in Belarusian Russian

Olga Goritskaya, Alexandra Chudar

Department for General Linguistics, Minsk State Linguistic University, Belarus

DOI:

## Abstract

The paper deals with differences in distributional patterns of lexical variants in Belarusian Russian. The study, based on the data collected through an online questionnaire, has shown that Belarusian Russian is characterised by quite rich lexical variation - both national (country-specific variants) and regional (items specific of one area of the country). The comparative study of onomasiological profiles of different age groups showed several instances of lexical change. Besides, various methodological questions, i.a. combination of corpus and experimental methods in sociolinguistic studies, are discussed. Questionnaires are considered not only as sources of data, but also as a way to interact with society.

Key words: lexical variation, Belarusian Russian, pluricentric language, sociolinguistics.

## Object of study

Belarus has two official languages: Belarusian, which primarily serves as an ethnic symbol (its use in everyday communication is limited), and Russian, which is a neutral means of communication. Besides, Belarusian-Russian mixed speech (Trasyanka) is used (Hentschel 2017). The Belarusian variety of the Russian language (Belarusian Russian) has its specific features at all linguistic levels. However, native speakers are partially aware of these features, which leads to frequent discussions on the “specificity” of Belarusian Russian as a separate language variety (Goritskaya 2018). The situation is complicated by the lack of empirical studies of the variety.

The aim of the present paper is to find the distributional patterns of lexical variants in Belarusian Russian, based on the data collected through an online questionnaire.

## Methodology

Recently, text corpora have become one of the main sources for studying lexical variation. However, our research on the Belarusian variety of Russian demonstrates that existing corpora have some limitations. Among the main difficulties of corpus-based research are studies of polysemous lexical variants and homonyms, low-frequency items, etc. Moreover, corpora provide data on word frequency only, with no information on its referent.

Besides, annotation of some corpora is not detailed enough, which results in lack of sociodemographic information on authors of texts, etc.

To complement corpus analysis, experimental methods, inter alia, questionnaires are used. We should take into account, though, that, contrary to corpus methods, questionnaires allow to investigate not the language itself, but the language from the speaker's perspective. However, the combination of several methods of analysis can provide more precise and reliable results.

The main part of our questionnaire consists of tasks where respondents are asked to name the objects in the pictures – everyday items (clothes, furniture, food, etc., cf. Iomdin 2014). Each question has a field for comments. Respondents can give more than one answer selecting words and expressions from the list and/or write down other variants in the field “other” or in the comments.

The questionnaire was spread through the internet. So far, we have collected about 1350 answers (1206 from Belarusian respondents, 90 from Russian ones, the remaining ones from other countries). To make sure that the variants are specific for Belarusian Russian, we have used a control group of native Russians (53 people).

## Results

To conduct our study, we have created onomasiological profiles (Geraerts 2010: 832) for 20 concepts, including synonyms and near-synonyms (not only words, but also multi-word expressions), and calculated their relative frequency.

The study has shown that the number of items in the onomasiological profiles can vary significantly – from 4 (*porečka*, *krasnaja smorodina*, *smorodina* and one occasional use of the word *kljukva* ‘cranberry’ for ‘red currant’) to 24 (*vatnik*, *vatovka*, *telogrejka*, *fufajka*, *kurtka*, etc. for ‘cotton wool-padded jacket’, including different expressive units).

In communication, country-specific variants compete with “universal” variants found in different countries, neologisms are used alongside archaisms, high-frequency words coexist with low-frequency ones, loanwords (from Belarusian, English etc.) compete with original items, which demonstrates quite rich lexical variation.

In our quantitative analysis, we have obtained the following results. First, we defined the proportion of variants specific for Belarusian Russian within the onomasiological profile. The per cent of variants characteristic of Belarus fluctuates between 0.93% (*berik/bèrak/bèrki* ‘a children’s play’) and 84.17% (*zakatki* ‘pickles’) (median – 29.48%).

Second, we identified lexical items typical of Belarus in general and specific of one area of the country. The comparison of onomasiological

profiles of native inhabitants and all residents shows the influence of migration on the language portrait of the modern city, therefore, to find out more about regional variation within the national, we considered the place where respondents spent their childhood, and not their current location. For instance, we have defined regions of Belarus with prevailing *stërka* (fixed in Russian dictionaries) or *stirka* (not found in Russian dictionaries, characteristic of Minsk and several other cities only) variant for ‘eraser’.

Naturally, significant variation can be observed in the names of children’s plays (18 words within the onomasiological profile). Our study has revealed that *dogonjalki* dominates in all Belarusian cities, though there are several variants specific for a certain location (*kvač*, *kič*, *bèrak/bèrik/bèrki*, etc.).

Third, the comparison of onomasiological profiles of respondents from different age groups allowed us to make some assumptions on Belarusian Russian language change. Some variants specific for Belarus decrease in their frequency over time: for example, *štroksy* ‘corduroy pants’ (the descriptive multi-word expression *vel’vetovye štany* ‘corduroy pants is gaining its popularity), *gol’f* ‘turtle neck sweater’ (the percent of a more “universal” *vodolazka* is increasing), etc. And vice versa, there are variants, the frequency of which is rising. This, inter alia, is relevant for *šufljedka* ‘drawer’ – one of the most salient markers of Belarusian Russian.

The qualitative analysis of comments points to the social meaning of lexical variation and provides an opportunity to study sociolinguistic awareness of speakers, as the respondents explicitly mention their associations with particular variants, share their emotions connected to words, discuss the “correctness” of lexical items, etc. Comments include data on peculiarities of usage of variants, associated with certain region, age, style of communication and other factors; influence of accommodation on usage of lexical variants, etc. Thus, metalinguistic comments supplement and/or confirm the quantitative data.

It is important to note that for a sociolinguist online questionnaire is not only a source of data, but also a way to interact with the society. For instance, in our study, several respondents expressed their gratitude for the opportunity to be heard and to have a look at their speech from outside.

## Discussion and conclusions

Despite the fact that the specificity of Belarusian Russian is often questioned, the findings of this study prove that, at least at the lexical level, this variety has its distinctive features that require further consideration.

First of all, within the onomasiological profiles, country-specific variants with various usage frequency are found, which points to differences in their significance: there are nuclear items used regularly as well as occasional nominations. The number of items in the onomasiological profiles also

varies. Within the nucleus, it is quite similar for all concepts (2–4 words); differences are observed on the periphery that includes items with restricted usage.

Belarusian Russian is rather homogeneous. However, this study allowed to reveal several instances of regional variation within the national one and to refine the distribution area of some lexical variants. The comparison of onomasiological profiles from respondents of different ages showed that there are both country-specific variants, which usage is shrinking and increasing over time.

As for the methodological aspect of this study, online questionnaires proved to be a valuable tool for sociolinguistic analysis. On the one hand, they provide (to some extent subjective) information on functioning of lexical variants in a language variety. On the other hand, such analytical instruments can influence the respondents, in particular, raise their level of sociolinguistic awareness. Awareness of structural features of a particular language variety is one of the key concepts in sociolinguistics (Babel 2016), and this is of great importance for Belarusian Russian, as this variety is at the stage of its development.

### **Acknowledgements**

The study was supported by the grant of the Belarusian Republican Foundation for Fundamental Research (project G18M-062).

### **References**

- Babel, A. M. (ed.) 2016. *Awareness and Control in Sociolinguistic Research*. Cambridge. Cambridge University Press.
- Geeraerts, D. 2010. Lexical variation in space. In Auer, P., Schmidt, P. A. (eds.), *Language in space. An international handbook of linguistic variation. Volume 1: Theories and methods*, 821-837. Berlin/New York. Mouton de Gruyter.
- Goritskaya, O. 2018. Discussions about Belarusian Russian: Linguistic Units and Cultural Models. In Liashchova, L. (ed.), *The Explicit and the Implicit in Language and Speech*, 190-222. Newcastle upon Tyne. Cambridge Scholars Publishing.
- Hentschel, G. 2017. Eleven questions and answers about Belarusian-Russian Mixed Speech ('Trasjanka'). *Russian Linguistics* 41(1), 17-42.
- Iomdin, B. 2014. Naivnye predstavlenija o značenijax slov v russkom jazyke. *Antropologičeskij forum* 21, 82-92.

# Vocabulary internationalisms and start-ups

Agnieszka Grażul-Luft

Faculty of Humanities and Social Sciences, The State University of Applied Sciences in Płock, Poland

DOI:

## Abstract

The environment of start-ups functions both in the area of the English language and the national languages which borrow numerous elements from the former one. The analysis of the internationalisms adapted in the Polish language proves that the vocabulary takes on the new shades of meaning as well as evaluative (validation) elements. It means that international words, highly popular in the globalisation era, have only apparent meaning even in specialist lexicon since they are not devoid of the influences of language and culture of other areas. That, in turn, can exert influence on the effectiveness of international communication.

Key words: internationalisms, borrowings, specialist lexicon.

## The dynamics of globalisation vs. linguistic changes

The dynamics of globalisation is the phenomenon whose repercussions encompass the language sphere which leads to the unification of languages. We are currently dealing with the strong expansion of the English language and there is a lot indicating that the world will be dominated by the model in which people will be in a way bilingual making use of English and the national languages (Zuckermann 2003), hence, in consequence, the national languages will be doomed to specific destruction (Steiner 1975). The dynamic civilisation changes necessitate searching for the new methods of linguistic observation which will soon enable identifying the phenomena that may appear ephemeral yet conducive to the language systems.

The globalisation processes focus on the area of the activity of start-ups. That specific and highly popular area of business activity stemming from Silicon Valley in a natural way derives its professional vocabulary from the English language. The environment of start-ups – innovative, creative, dynamic – make simultaneous use of the English language, common and facilitating communication for all. On the other hand, it functions in the local language area where there are inevitably new English elements introduced adapted to the local language systems. That kind of adaptation is often connected with the diversification of meanings of the lexemes creating some language shades and validation elements (emotional load) absent from other languages.

Start-ups are a new form of business activity so the language of that specific environment has not yet constituted the subject of detailed analysis. Undertaking its research one has to face two main challenges: globalism (the analysis has to be rooted in the wide context) and the dynamics as well as the liability of functioning (in separate language areas) which may be compared to the brisk stream. The research was conducted by means of the Internet network which was treated as a scattered corpus (Kilgariff 2003) (using the tools which include search engines).

The analysis was carried out in the area of the Polish language yet, it seems that the conclusions drawn illustrating the element of the globalisation processes may be referred to not only other languages but also to other areas of human activity.

### **Internationalisms amongst start-ups**

In the case of the vocabulary connected with the activity of start-ups the borrowed professionalisms are internationalisms i.e., according to the classic definition of Haugen, the vocabulary of similar meaning and spelling with often distinct pronunciation spread in the majority of languages of western Europe (Haugen 1950). The vocabulary analysed functions not only in European languages but probably in all languages used in the areas where start-ups operate.

### **Methods**

In the research presented there were 12 internationalisms analysed connected with the activity of start-ups which function in the area of the Polish language: *akcelerator* (Eng. *accelerator*), *bootstrapping* (Eng. *bootstrapping*), *decacorn* (Eng. *decacorn*), *ekosystem* (Eng. *ecosystem*), *inkubator* (Eng. *incubator*), *jednorożec* (Eng. *unicorn*), *lean startup* (Eng. *lean startup*), *MVP* (Eng. *MVP*), *pitch deck* (Eng. *pitch deck*), *pivot* (Eng. *pivot*), *skalowanie* (Eng. *scaling*), *start-up/startup* (Eng. *start-up/startup*).

The research was multi-stage, it checked whether the given internationalism appears, how it is defined, in what meanings it is used in the dictionaries of the English and Polish languages, in the National Corpus of Polish, specialist papers, (the original version, translations and Polish-language ones) and in non-specialist texts (popular press, Internet websites, blogs etc.).

The vocabulary connected with that type of activity is in its early stages of dissemination and, in most cases, is not yet noted in the dictionaries. The analysis conducted was qualitative, pragmalinguistic, which enabled making comparison and identifying the similarities and differences in the meaning of the internationalisms in the areas of the English and the Polish languages.

### **Examples of diversifying of the meanings**

Among the 12 analysed internationalisms one can indicate both those in case of which there were no transformations of meaning in the process of moving to the area of the Polish language and those in which there were some deviations or diversifications stemming from separate mechanisms.

#### **1. Changes in meaning stemming from translation e.g.**

The acronym: *MVP* - Minimum Viable Product; “smallest group of features that will elicit customer feedback” (Blank, Dorf 2012). In the Polish translation of the book by Blank and Dorf *MVP* is „minimalnie satysfakcjonujący produkt” (Blank, Dorf 2013), The word *satysfakcjonujący*, closer to English *satisfying*, does not seem right in that context. Who and in what way would the product satisfy (its creator, the customer or the investor)? *Viable* has in its meaning range: *profitable, likely to succeed, sustain on the market, competitive*. Experienced start-uppers explain that they mean the product ready enough to be demonstrated to the customer or the investor to assess the interest (with the lowest possible financial expenses and asap). That translation appeared in the start-up “Bible” hence it is constantly reiterated.

#### **2. Emotional load**

The start-ups are currently highly popular and strongly promoted in Poland. It is trendy not only to possess a company but rather a start-up. That causes fading of the meaning of start-up originally specifying the activity, but also adding the internationalism from the area of start-up activity of the emotional load – in that case a positive one. An interesting example is the internationalism *pitch deck*, which denotes the presentation for the investor. The appearance of that word in the Polish language instead of the common *prezentacja*, proves the need for exemplification, emphasis and sophistication. Pitch deck is a presentation “well-thought over”, „effective”, „catchy”, while the ability to create it is „improved throughout start-up life”.

#### **3. Metaphorisation and the perception of the World**

Another example is the internationalism *unicorn* denoting the start-up worth at least a billion dollars. The creator of the metaphor, an American investor, wanted to emphasize that a start-up worth a billion dollars is as rare as a unicorn – a mythological creature. The rarity of its appearance – the construction block of the meaning of the lexeme *unicorn* in the area of the American language is not yet found in the Polish language. In the Polish cultural sphere a unicorn is not regarded as a real creature thus the metaphor entails rather an unachievable dream.

## Conclusions

The internationalisation of the vocabulary is somehow accompanied by a disharmony since the language speakers employ the same vocabulary yet the understanding of the terms used may be slightly different. The analysis enabled not only researching the problem in the new area but also reflecting on the globalisation phenomena which are becoming the key to the nations functioning in many a sphere such as in culture, business and language. Realizing such mechanisms is essential in the perspective of international communication which, despite the increasing popularity of the English language, is not devoid of the influence of the national and cultural features - even in the case of internationalisms.

The analysis was conducted within the project „Language as the tool serving the increase in the effectiveness of the activity of start-ups – pragmalinguistic research”, realized within the grant financed out of the resources of the Commune-Town of Płock.

## References

- Blank, S., Dorf, B. 2012. *The Startup Owner's Manual: The Step-by-Step Guide for Building a Great Company*. California: K&S Ranch Inc.
- Blank, S., Dorf, B. 2013. *Podręcznik startupu. Budowa wielkiej firmy krok po kroku*. Gliwice: Helion S.A.
- Haugen, E. 1950. The Analysis of linguistic borrowing. *Language* 26, 210-231.
- Kilgarriff, A., Grefenstette, G. 2003. Web as Corpus. *Computational Linguistics* 29 (3), 333-347.
- Steiner, G. 1975. *After Babel: Aspects of language and translation*. Oxford University Press.
- Zuckermann, G. 2003. Language contact and globalisation: The camouflaged influence of English on the world's languages—with special attention to Israeli (sic) and Mandarin. *Cambridge Review of International Affairs* 16 (2), 287-307.

# Experimental linguistics in aphasia and child language

Maria Irma Hadler Coudry, Isabella de Cássia Netto Moutinho  
Department of Linguistics, State University of Campinas, Brazil  
DOI:

## Abstract

The discursive approach of language studies in aphasia and childhood is the theme of this text and derives from classical studies (Jakobson, 1941). Such perspective is based on the theoretical-methodological assumptions postulated by Discursive Neurolinguistics, which aggregates authors who study language and brain in action, its functioning, normal and pathological. Experimental activities (which do not include standardized tests) are analysed, in different discursive genres, produced along with longitudinal studies of aphasics and children in (re)acquisition of speech and writing. Such data happen in different discursive situations, the way language works in our society. Language is seen as work and action between interlocutors, incomplete, heterogeneous and polysemic. The data confirm the classic hypothesis of the relation between aphasia and child language.

Key words: language, aphasia, children, speech, writing.

## Language, brain and subject

Our language conception is postulated by Franchi (2002), who conceives it as a constitutive activity and as a work, a historical product of the relationship between speakers, which meaning depends on a set of aspects, not put beforehand, but conditioned on who speaks, with whom, under certain circumstances, wondering what, etc. Thus, we reinforce the creative possibility of language, the force that moves saying, repeating, meaning. From this conception derives a subject that deals with language, captured by it (De Lemos, 2002).

An incomplete subject, in need of the other and determined by devices (Foucault, 1969; Agamben, 2009) that define actions, language and gestures. A subject who works the language, in interactions with his others, even as aphasic, even as a child. The conception of the brain we take is the one from Soviet Psychology, (Vygotsky, 1926; Luria, 1981), which conceives it from a historical point of view, the product of human action inserted in a culture and shared by a society.

A hierarchically functioning brain that articulates its areas and zones for the complexity of the activity the subject performs. According to Luria (1981), there are definite roles in terms of relationships between brain areas which derive from learning and use (automated and creative) that rely on

neuroplasticity, inherent to an incomplete system and subject to new configurations.

## Methodology

The methodology we use in ND is heuristic, based on discovery procedures that the researcher, theoretically oriented, proposes as activities with language. Once it is discursively oriented, not based on standard tests (Coudry 1986, 2018), which activities are always the same for different subjects, our activities may vary from subject to subject and the solutions they find may also differ. It is important, in the interaction with the subject, that he understands what is required of him, that is, that what is proposed makes sense to him. This methodology is not ready (like a test), but the theoretical foundations that support it are so established and solid that the researcher proposes and adjusts the activities for each subject. Interestingly, it is a process of discovery that feeds the theoretical researches that, in turn, expand the methodological possibilities. What is experimental about our approach is the possibility of proposing experiments with the language to help the subject, that is, that contribute to the researcher and to the subject to deepen their knowledge about the difficulties and the possible solutions. Language experiments for aphasics and children<sup>1</sup> are conducted for the purposes of understanding both the difficulties and possibilities for a problem to be solved (solving problems: hypothesis, execution, verification). Writing: similar processes

This first experiment relates a child trying to fill in a crossword– the word *gema*, which means egg yolk. MT is a child that has attended Ccazinho because of his school difficulties. He is at the age of 12, at the sixth grade and still hasn't learnt to read and write. First, MT writes *gma* and realizes that something is wrong, because there is a crossword field left to fill in. He points his finger at the letter G and reads (g), then points the letter E, and reads (ma). The letter G, because its name coincides with the sound of the syllable to be written, leads MT to write only the consonant, without spelling out the core of the syllable. The investigator's intervention is that MT disentangles himself from the name of the letter to write the syllable GE (Coudry 2010, Coudry Bordin 2012, Barthelson 2014). We present now the data of an aphasic subject, RS, who goes through the same process in his reentry in writing (Coudry, Freire, Gomes 2006). In writing his mother's name (Rosangela) omitted the syllable core, because the vowel (E) is in the letter name (G). He writes ROSANG LA and leaves a blank space between the letters (G) and (L).

In order to write the full mother's name and to write the word *gema*, both the child and the aphasic must erase the sound image of the letter G to be able to write the letter E. In the child's case, the perception that a second

writing is needed is the activity itself, which is the crossword – there is an empty place to be filled. In the case of the aphasic, the reading of what he wrote and the mark of the blank space were fundamental for him to understand the adjustment he would need to make.

The Discursive Neurolinguistics researches has also found out other kinds of similarities through other different kinds of language experiments:

In phonology, entering the alphabetic system, using speech, relating sound and letter; in prosody - when the response intonation does not match to its content but repeats the speaker's words. In morphology, when derivation processes act in the formation of words: child says *arrumeira* and not *arrumadeira* (maid) and aphasics say *gentoas* and not *gente* (people). In syntax: the aphasic says *menino trem* (boy train) and child the child says *Pedro duiche* (*Pedro* (sand)*wich*) which shows that they preferentially say names to express the action between name and verb. In semantics, when the child seeks meanings of new words and aphasic produces paraphasias (*gentoas*) as well as metonymic descriptions (*that which shows our legs* – trying to define *bermuda*). In discourse: when the reversibility of discursive roles is done by repeating the other's speech Question: *Do you go?* Answer: *You go.*), which shows no personal implications. All in all, the longitudinal follow up of aphasics and children, we can find data (Courdry 1996) - which illuminates the speaking, writing and reading processes and exposes the discursive event, moving theorizing and advancing the analysis.

## Conclusions

The data confirm the classic hypothesis about the interrelationship of aphasia with childhood language, strongly indicating that the transformations that aphasia causes in speech, reading and writing resume childhood language processes, or that incomplete childhood processes occur again in aphasia. There is, therefore, much similarity between processes that occur in language changes in aphasia, predicted in the language itself, that co-occur in language acquisition, that is, in the two poles of human life. The ND approach brings discursive practice as a methodological resource - in which experiments with language can be included - which, unlike tests, not only points the child's and aphasic limits with speech, reading and writing, but also possibilities of intervention, which are different from the traditional clinical and pedagogical approach. which disregard the various faces of the subject of language: historicity, psycho-affectivity and singularity. Thus, the inevitable encounter between aphasia and childhood offers an approach that revolutionizes the old approaches to language assessment and intervention.

## Notes

The functioning of these Centers (the first one aimed at aphasics and the second one for children and young people with learning difficulties, especially reading and writing, and speech is also considered) is based on verbal and nonverbal discursive practices that make sense to participants who live together. themselves systematically and who recognize themselves as a group of people and as speakers of a natural language for sharing, with differences, cultural parameters (history, habits, values, beliefs, artefacts, devices, etc.) that guide what is spoken, write, read.

## References

- Aganbem, G. 2009. O que é contemporâneo? E outros ensaios. Chapecó: Argos.
- Barthelson, B. 2014. Relações de sentido no processo de aquisição e uso da leitura e da escrita. Dissertação de mestrado. Campinas, IEL/Unicamp.
- Coudry, M. 1986. Diário de Narciso: discurso e afasia. Tese de Doutorado. Campinas, IEL/Unicamp.
- Coudry, M. O que é dado em Neurolinguística? In: CASTRO, M. F. P. (Org.), Método e o Dado no Estudo da Linguagem. Campinas: Editora da Unicamp, 1991/1996, pp. 179-194.
- Coudry, M. 2018. Diário de Narciso e Neurolinguística Discursiva: 30 anos depois. Cadernos De Estudos Lingüísticos, 60(2), 323-350.
- Coudry, M; Freire, F.; Gomes, T. 2006. Sem falar, escrever e ler e ainda sujeito da linguagem. Estudos Lingüísticos XXXV, p. 1375-1384. Grupo de Estudos Linguísticos, Brasil.
- Coudry, M., Bordin, S. 2012. Afasia e infância: registro do (in)esquecível. Cadernos de Estudos Linguísticos, 54 (1) Campinas, UNICAMP.
- Coudry, M. 2010. Relatório de Pesquisa: Projeto Integrado em Neurolinguística: avaliação e banco de dados. CNPq/301726/2006-0, 65p. Impresso. 2010.
- De Lemos, C. 2002. Das vicissitudes da fala da criança e de sua investigação. Cadernos de Estudos Linguísticos, p. 41-69. Campinas, IEL/Unicamp.
- Foucault, M. Arqueologia do Saber. Petrópolis Lisboa: Vozes, 1969.
- Franchi, C. 2002. Linguagem: atividade constitutiva. Cadernos de Estudos Linguísticos, n° especial, 37-74.
- Jakobson, R. 1941. Langage enfantin et aphasie. 1ère. Edition. Paris, Flammarion.
- Luria, A. 1981. Fundamentos de Neuropsicologia. 1a ed. São Paulo, EDUSP.
- Vygotsky, L. 2004. Psicologia Pedagógica. São Paulo, Martins Fontes.

# Lexical diversity and Mild Cognitive Impairment

Sofie Johansson<sup>1</sup>, Kristina Lundholm Fors<sup>1</sup>, Malin Antonsson<sup>1</sup>, Dimitrios Kokkinakis<sup>1,2</sup>

<sup>1</sup>Department of Swedish, University of Gothenburg, Sweden

<sup>2</sup>Centre for Ageing and Health, University of Gothenburg, Sweden

DOI:

## Abstract

This paper explores the role that various lexical-based measures play for differentiating between individuals with mild forms of cognitive impairment (MCI) and healthy controls (HC). Recent research underscores the importance of language and linguistic analysis as essential components that can contribute to a variety of sensitive cognitive measures for the identification of milder forms of cognitive impairment. Subtle language changes serve as a sign that an individual's cognitive functions have been impacted, potentially leading to early diagnosis. Our research aims to identify linguistic biomarkers that could distinguish between individuals with MCI and HC and also be useful in predicting MCI.

Key words: mild cognitive impairment/MCI; lexical diversity; language; Swedish

## Introduction

The number of people living with dementia worldwide is projected to be 65.7 million in 2030, and this number will double every 20 years *cf.* Prince et al. (2013). It has been established that early diagnosis is beneficial and ultimately current research is exploring the possibility of identifying persons with mild forms of cognitive impairment at an early stage.

Mild cognitive impairment (MCI) is a condition characterized by cognitive decline greater than expected for an individual's age and education level. As the MCI progresses, MCI individuals face a higher risk of developing Alzheimer's Disease (AD). While language impairments have been well described in AD, language impairment in people with MCI is less well understood and the need for further research in all aspects of language and during all stages of the disease, has been recently emphasized in various studies; *cf.* Laske et al. 2015, Boschi et al. 2017; Beltrami et al. 2018.

In this paper, we examine lexical aspects of lexical diversity of individuals with MCI and age-matched controls and apply various measures described in relevant literature trying to identify whether there are correlations between lexical diversity measures and the participants' characteristics. The objective of the current study is to determine which and to what extent various lexical diversity measures can differentiate the two

groups and what is the range of scores of the measures found in transcriptions of the oral narratives.

## Background

Multiple components of language can be assessed in order to reveal linguistic features that are likely to serve as discriminators between individuals with MCI and cognitively healthy controls. One of these components is at the lexical level, but research so far has shown conflicting results, probably because impairment may only appear at the *semantic* and *macrolinguistic* language levels on e.g. discourse relations (Masrani et al., 2017). Aramaki et al. (2016) used transcriptions of both written and spoken samples in order to measure various language ability scores. Their analysis of the spoken narrative showed e.g. that MCIs had a significantly larger vocabulary size which might indicate compensatory behavior for MCI persons. Fergadiotis et al. (2013) collected validity evidence regarding techniques for measuring lexical diversity for the study of aphasic discourse. Two of the tested lexical diversity scores, the Measure of Textual Lexical Diversity, MTLTD (McCarthy, Jarvis, 2010) and the Moving-Average Type-Token Ratio, MATTR (Covington, McFall, 2010) yielded the strongest evidence for producing unbiased lexical diversity scores, suggesting that they may be the best measures for lexical diversity in people with aphasia. For other relevant research on lexical profiles and linguistic features cf. Laufer & Nation (1995) and Biber (1995).

## Data, methods, results and limitations

Participants for this study were recruited from the longitudinal Gothenburg MCI study (Wallin et al., 2016). All subjects were native speakers of Swedish, and all studies are approved by the local ethical committee review board. In the current analyses, we include only participants who had completed the Cookie Theft task (Goodglass et al., 1983), a widely used test to elicit narrative speech. Participants were asked to describe everything they saw in the picture, and to talk for as long as they liked. The narratives were digitally recorded and manually transcribed according to a detailed protocol developed by the authors. Only the orthographic transcriptions of the spoken samples for each participant were used in the measurement of lexical diversity in this study. For each participant, we applied 16 lexical measures, e.g. long words, hapax legomena, frequency-related coverage at a general and a more specific level, originality, erroneous or non-existent words, contextual and non-contextual descriptors. As indicated by e.g. Laufer, Nation (1995), Richards et al. (2009) and Johansson, Ohlsson (2019), there are some lexical features which it is possible to relate to prominent or more proficient oral or written language, e.g. long words, hapax legomena and

originality. The specific lexical measures in this work were selected since a larger vocabulary and richer language is commonly indicated by a higher lexical diversity, and contextual, rather than general language and vocabulary size, *cf.* Nation (2013) and Milton (2009). We assumed that there would be individual variables co-occurring as well as variables which would be significant to a particular group of individuals in the study. In order to find possible correlations between the lexical measures, a bivariate 2-tailed correlation using Pearson's coefficient was conducted using IBM SPSS v 25.

The length of speech varied among the participants; the transcribed spoken picture descriptions ranged from 57 to 617 tokens (mean 185.47) for the HCs; and 46 to 481 tokens (mean 185.45) for the MCIs. The correlations which could be related to individual differences or similarities were found to be age-related to common words (.395\*\*) and the 1000 most common tokens used in a Swedish everyday context (.435\*\*) and negative correlation to long words and contextual descriptors. Other individual correlations were contextual descriptors which were related to hapax legomena (.439\*\*) and long words (.439\*\*). Weak correlations related to the two groups (HC, MCI), were contextual descriptors (.267\*) regarding the HC group, and erroneous words (.259\*) regarding the MCIs. This study has limitations. First, our study comprised relatively few participants, 29 HCs and 26 MCI. As this was not a very large sample, future studies with larger sample sizes are needed to verify our findings. Next, the lexical analysis carried out to examine differences in elicited verbal production originates from research in language development and language proficiency which might affect the generalizability of the results, although these measures have been applied to communication disorders such as aphasia. Preliminary findings indicate that there is a difference in how the individuals express themselves which might be related to age. It seems that older participants use more common words as opposed to more frequent use of contextual descriptors. This difference seems to be unrelated to numbers of education years. Further studies would involve a more sophisticated categorization of linguistically descriptive features, such as active-passive use of verbs, and also, lexical density where the number of content words is compared to functions words.

### **Conclusions and future work**

Our preliminary results show some differences between the scores and groups. The validity needs to be further investigated using more lexically-based measures of lexical diversity and richness in order to capture the degree of difference between the groups, by applying e.g. clustering. Further investigations are also required to shed light on the relationship between the lexical diversity measures and neuropsychological tests, and also apply them to lemmatized transcriptions. In the near future, we also plan to augment

these measures e.g. MTLT, MATTR, and also compare the measures at 2 points in time, since exactly the same cohort has recently repeated the same task and have also undergone renewed neuropsychological assessments.

## Acknowledgements

This work has received support from *Riksbankens Jubileumsfond* – the Swedish Foundation for Humanities & Social Sciences; grant NHS 14-1761:1.

## References

- Aramaki, E, Shikata, S, Miyabe, M., Kinoshita, A. 2016. Vocabulary size in speech may be an early indicator of cognitive impairment. *PLOS ONE* 11(5).
- Beltrami, D., Gagliardi, G., Rossini Favretti, R., Ghidoni, E., Tamburini, F., Calzà, L. 2018. Speech analysis by NLP techniques: A possible tool for very early detection of cognitive decline? *Frontiers in Aging Neuroscience*. 10.
- Biber, D. 1995. *Dimensions of Register Variation: A Cross-Ling Comparison*. CUP.
- Boschi, V., Catricalà E., Consonni M., Chesi C., Moro A., Cappa S.F. 2017. Connected speech in neurodegenerative language disorders: a review. *Frontiers in Psychology*, 8:269.
- Covington, M., McFall, J. 2010. Cutting the gordian knot: The moving-average type token ratio (MATTR). *J. of Quan. Ling.* 17, 94-100.
- Fergadiotis, G. Wright, HH., Westa, TM. 2013. Measuring lexical diversity in narrative discourse of people with aphasia. *Am J Speech Lang Pathol.* 22, 2.
- Goodglass, H. Barresi, B., Kaplan, E. 1983. *Boston Diagn Aphasia Exam*. Kluwer.
- Johansson, S., Ohlsson, E. 2019. Visualizing vocabulary. *Multiling Matters*.
- Laske, C. et al. 2015. Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer's & Dementia* 11(5), 561-578.
- Laufer B., Nation, P. 1995. Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics*. 16:3. pp. 307-322.
- Masrani, V., Murray, G., Shoshana Field, T., Carenini, G. 2017. Domain Adaptation for Detecting MCI. *Advances in AI*. Pp 248-259. Springer.
- McCarthy, P., Jarvis, S. 2010. MTLT, voc-d, and HD-d: *Beh Res Meth.* 42, 381-392.
- Milton, J. 2009. *Measuring Second Language Vocab Acquisition*. *Multiling Matters*.
- Nation, ISP. 2013. *Learning Vocabulary in Another Language*. *Cambr. Appl. Ling.*
- Prince, M. et al. 2013. The global prevalence of dementia. *Alz & Dem* 9(1), 63-75.
- Richards B., Malvern, D.D., Meara, P., Milton, J., Treffers-Daller, J. 2009. *Vocabulary Studies in 1st and 2nd Language Acquisition. The Interface Between Theory and Application*. Palgrave Macmillan.
- Wallin A., Nordlund A., Jonsson M., Lind K., Edman Å., Göthlin M., Stålhammar J., Eckerström M., Kern S., Börjesson-Hanson A., Carlsson M., Olsson E., Zetterberg H., Blennow K., Svensson J., Öhrfelt A., Bjerke M., Rolstad S., Eckerström C. The Gothenburg mci study. *J of Cerebral Blood Flow and Metabolism*, 36(1):114–31.

# Linguistic Intraoperative Protocol

Luis M.T. Jesus<sup>1,2</sup>, Joana Alves<sup>1</sup>, Mafalda Cardoso<sup>1</sup>, Mariana Morgado<sup>1</sup>

<sup>1</sup>School of Health Sciences (ESSUA), University of Aveiro, Portugal

<sup>2</sup>Institute of Electronics and Informatics Engineering of Aveiro (IEETA), University of Aveiro, Portugal

DOI:

## Abstract

The Dutch Linguistic Intraoperative Protocol – EP (DuLIP-EP) is a culturally adapted and validated instrument that allows language evaluation in patients with low-grade gliomas submitted to awake brain surgery. It assesses language areas which are related to different brain regions and pathways. In this paper, each task included in DuLIP-EP is presented and some preliminary results from its application to a normal Portuguese population are discussed.

Key words: language, brain, mapping, awake surgery

## Introduction

The Dutch Linguistic Intraoperative Protocol (DuLIP) was originally developed in the Netherlands, to evaluate pre-, intra-, and post-operative language skills in patients with low-grade gliomas (De Witte et al., 2015). In Portugal, there are no validated instruments to assess language skills during surgery. This study aims to translate to European Portuguese (EP), culturally adapt and validate the DuLIP in the general population. This protocol includes phonological, semantics, syntax and articulation tasks, that are applied to the patients according to their tumour localisation and its functional implications.

## Methods and description of the tests

The DuLIP-EP was initially based on a literal translation of the original tool, but the resulting materials were also adapted according to variables such as frequency, imageability, age of acquisition, prevalence and word class. Specific criteria, such as complexity levels, number of phonemes, number of syllables and syllabic structure, were used to build the final version of the repetition task. Furthermore, in the tasks that involve phrases, criteria such as time/verbal mode, order of constituents in the sentence, number of words, passive/active voice, were used.

To validate the final version of the protocol, test administration, registration and correction of the items were discussed by the three authors of this paper, particularly when uncertainties regarding the scoring, until a consensus was found.

Seventy (70) volunteers were randomly selected from a larger database for analysis (data collection still underway). Inclusion criteria were defined according to the original study (De Witte et al., 2015), but considering that native language should be EP.

Participants were 25 male and 45 females. Regarding age, two groups were created:  $\geq 55$  years old ( $n=12$ ) and  $\geq 18 < 55$  ( $n=68$ ). Ages varied between 18 and 89, with an average (AVG) of 37.20 [standard deviation (SD) = 16.06]. Fourteen (14) participants had less than 12 years of schooling (AVG=15.03; SD=4.10). Addenbrooke's Cognitive Examination scores varied between 76 and 100 (AVG = 15.03; SD = 4,64).

In all tasks performed there is an intervention of various brain areas, so each assignment evaluates and stimulates different areas. Moreover, some tasks have a time frame so that Direct Electrical Stimulation (DES) can be used to assess the function role of restricted cortical and subcortical regions while in surgery.

### **Phonology**

Phonology tasks include: Repetition of words (4s time frame) – This task aims to evaluate the ability to produce words according to their auditory discrimination. It contemplates 300 stimuli organised by different levels of complexity; Phonological odd word out (without time frame) – The objective of this task is to evaluate the phonological awareness. There are 15 lists with 4 monosyllabic words where participants should identify the intruder; Phonological sentence judgment – This task contemplates 30 sentences, which must be validated by the participants as being correct or incorrect. The incorrect ones are those containing non-words; Phonological fluency – This task evaluates the capacity of phonological evocation. In a minute, the participants should say the maximum number of words beginning with /m/, /p/ and /r/ phonemes.

### **Semantics**

Testing of semantics is based on: Reading with semantic odd word out (4s time frame) – The participants must read aloud the word that doesn't fit semantically a list of 4 words (25 items); Reading with semantic odd picture out – Participants have to name the picture that doesn't fit semantically in a group of 3 drawings (25 items); Semantic association

Participants should read two words and add a third semantically linked to the previous ones (25 items); Sentence completion (close context, 4s time frame) – Two tasks to evaluate this capacity. One is prompted by a specific context, the other evaluates speech production with an open context/less semantically induced. Participants must read out aloud and complete sentences in a meaningful way (25 items per task); Semantic sentence

judgment (without time frame) – The goal of this task is to discriminate between semantically correct and incorrect sentences (50 items); Semantic fluency – Within a minute each, participants have to name as many animals and jobs as possible.

### **Syntax**

Syntax evaluation tasks include: Verb generation (4s time frame) – Singular nouns are presented in order and participants are asked to semantically relate them to a verb, evoking them; Syntactic sentence judgment (without time frame) – This task aims to assess the capacity to discriminate between 50 correct and incorrect sentences, syntactically speaking; Syntactic fluency – During one minute, the subjects have to produce as many verbs as possible.

### **Articulation**

Verbal diadochokinesis: This task includes 9 sequences that the participants should produce as fast and accurate as possible. This task should be timed.

Naming (4s time frame)

Object naming of 100 black and white drawings.

### **Results**

Global phonology scores were: Repetition Task (AVG = 299.17; SD=1.063), in a total of 300 items; Phonological Intruder Analysis (AVG = 14.56; SD = 1.04), in a total of 15 items; Phonological Sentence Judgment (AVG = 29.96; SD = 0.20), in a total of 30 items; Phonological Fluency of /m/, /r/ and /p/ phonemes (AVG = 14.56; SD = 6.03), (AVG = 18.26; SD = 6.05) and (AVG = 18.90; SD = 6.86), respectively.

Semantics scores results were: Semantic Odd Word Out (AVG = 23.51; SD = 2.22) and Odd Picture Out (AVG = 24.33; SD = 1.19), out of 25 pictures and words; Semantic Association (AVG = 21.81; SD = 2.42) in a total of 25; Semantic Completion with close context (AVG = 24.50; SD = 1.09), lower scores when compared with the less induced context (AVG = 24.93; SD = 0.31), each one with 25 items. Semantic Sentence Judgment (AVG = 48.89; SD= 1.40), in a total of 50 items; Semantic Fluency: animals and jobs (AVG = 24.17; SD = 6.05) and (AVG = 20.06, SD = 5.24), respectively.

Syntactic tasks scores were: Verb Generation (AVG = 47.14; SD = 4.37) in a total of 50; Syntactic Sentence Judgement I & II (AVG = 24.54; SD = 0.63) and (AVG = 24.61; SD = 0.77), respectively, each with a total of 25 items; Syntactic fluency (AVG = 24.66; SD = 7.82).

Object naming of 100 pictures average scores were 96.37 with a standard deviation of 3.65.

Diadochokinesis scores for each task were: Task 1 – AVG = 3.03; (SD = 1.30); Task 2 – AVG = 2.07 (SD = 1.04); Task 3 – AVG = 2.78; (SD = 0.93); Task 4 – AVG = 3.25 (SD = 1.00) Task 5 – AVG = 5.50 (SD = 1.62); Task 6 – AVG = 7.44 (SD = 2.72); Task 7 – AVG = 6.06 (SD=2.38); Task 8 – AVG = 6.52 (SD=2.15); Task 9 – AVG = 6.82 (SD=2.19).

## Conclusions

In general, most of the tasks have a mean score close to the maximum score desired which could indicate that the tasks are well tailored and well perceived by the participants. Yet, some adjustments are needed. Since this was a preliminary test, the SD values were higher than desirable, but we aim to expand testing to 140 participants in order to reach more solid results. As future work it would be useful to apply the DuLIP-EP protocol to some clinical cases as well study possible correlations between variables.

## Acknowledgements

The authors would like to thank Evy Brink, Djaina Satoer and Elke Witte. This research was funded by the Foundation for Science and Technology (FCT), in the context of the project UID/CEC/00127/2019.

## References

- De Witte, E., Satoer, D., Robert, E., Colle, H., Verheyen, S., Visch-Brink, E., Mariën, P. 2015. The Dutch Linguistic Intraoperative Protocol: A valid linguistic approach to awake brain surgery. *Brain and Language*, 140, 35–48.

# **Animacy distinctions and semantic dementia**

Anna Kaglik<sup>1</sup>, Clara Sanches<sup>1</sup>, Elsa Mhanna<sup>1</sup>, Raffaella Migliaccio<sup>1,2</sup>, Marc Teichmann<sup>1,2</sup>, Antoni Valero-Cabré<sup>1</sup>

<sup>1</sup>Institut du Cerveau et la Moelle Épineière, Sorbonne Université, France.

<sup>2</sup>Department of Neurology, Assistance Publique Hôpitaux de Paris, France.

DOI:

## **Abstract**

We explored relationships between performance in “living” and “non-living” animacy distinctions and brain metabolism in patients with a semantic variant of a primary progressive aphasia (Sv-PPA) as well as healthy volunteers. Outcomes did not convincingly support differences in the processing of living’ vs ‘non-living’ items in Sv-PPA patients, and point out at complex potentially interesting anatomical specialization in the processing of these two types of categories, which should be further pursued in a larger cohort.

Key words: animacy distinction, impairment, semantic category, hypometabolism, temporal lobes, correlation.

## **Introduction**

The difficulty in word-finding is the most evocative of language disorder. This impairment is characteristic of sv-PPA, a progressive neurodegenerative disorder affecting semantic capacities. However, not all semantic concepts seem to be identically affected. Prior research has shown that the conceptual category of ‘living’ (divided into ‘animate’ and ‘non-animate’) is particularly affected, independently from other semantic categories. Nonetheless, this finding has not always been replicated and remains controversial. Sv-PPA patients show clear signs of atrophy and hypometabolism in anterior temporal lobes (ATLs) with a left side predominance. ATL plays a crucial role in semantics. The ATLs encompass a region composed of cortical areas, which differ significantly with regards to their cytoarchitectural organization and connectivity patterns (Kondo et al. 2003). Here, we aimed to explore: dissociations between ‘living’ vs ‘non-living’ items processing in Sv-PPA, using a correlational approach between behavioral outcomes in semantic tasks and regional metabolic maps in areas of temporal lobes.

## **Materials and methods**

Fourteen Sv-PPA patients and five healthy subjects (HS) were recruited. All measurements were obtained as part of an ongoing clinical trial sponsored by AP-HP (PHRC National ‘Stim-SD’). Analyses focused on three computer-

based language tasks: 1) a *verbal semantic association* task (AS verbal / 26 trials); 2) a *visual semantic association* task (AS visual / 26 trials); 3) a *categorical decision* task (CD task / 40 trials). The same day, participants underwent a PET-MRI study. Accuracy and reaction times (RT) of correct responses for ‘living’ and ‘non-living’ items were separately recorded. As well as Inverse Efficiency (IE) measure, calculated by dividing the RT by 1–proportion of Errors (PE) for each category in order to factor jointly *speed* and *accuracy* of correct responses during task execution. PET analyses employed the AAL atlas and focused on the six regions of interest: bilateral Middle Temporal Pole, Superior Temporal Pole and Parahippocampal Gyrus. Non-parametric Wilcoxon test was used to estimate the statistical significance between semantic categories, and Spearman or Pearson for correlation coefficients between metabolic levels and categorization performance.

## Results

Accuracy differences between the two item categories reached significant differences ( $p=0.03787$ ) only for the semantic categorization (CD) task, showing higher accuracy for ‘non-living’ compared with ‘living’ items. However, the RT for correct responses was faster for ‘living’ than for ‘non-living’ category ( $p = 0.004028$ ). A trend ( $p>0.05$ ) for IE suggests higher values for ‘living’ vs ‘non-living’ items. For HS, no statistical significance was found for accuracy, despite a benefit of correct responses for “non-living” category observed.

Most significant correlations in Sv-PPA patients concerned the processing of ‘non-living’ category (Accuracy and IE) and FDG metabolism in right temporal pole subregions. The correlations identified category-specific temporal lobe regions i) in the superior temporal pole for ‘living’ items, ii) and in the right middle temporal pole for ‘nonliving’ items. The right superior temporal pole metabolism scaled with the performance of both ‘non-living’ but also ‘living’ items. Additionally, bilateral parahippocampal cortex showed exquisite category- and task-specific dependency for ‘living’ items only when tested with the AS visual semantic task. Readers should bear in mind that none of the correlations presented in this section survived at the Bonferroni correction for multiple comparisons (corrected thresholds at  $p<0.008$ ). Hence at the current stage, the above-reported correlations should be considered significant using uncorrected statistical approaches, hence be taken very cautiously.

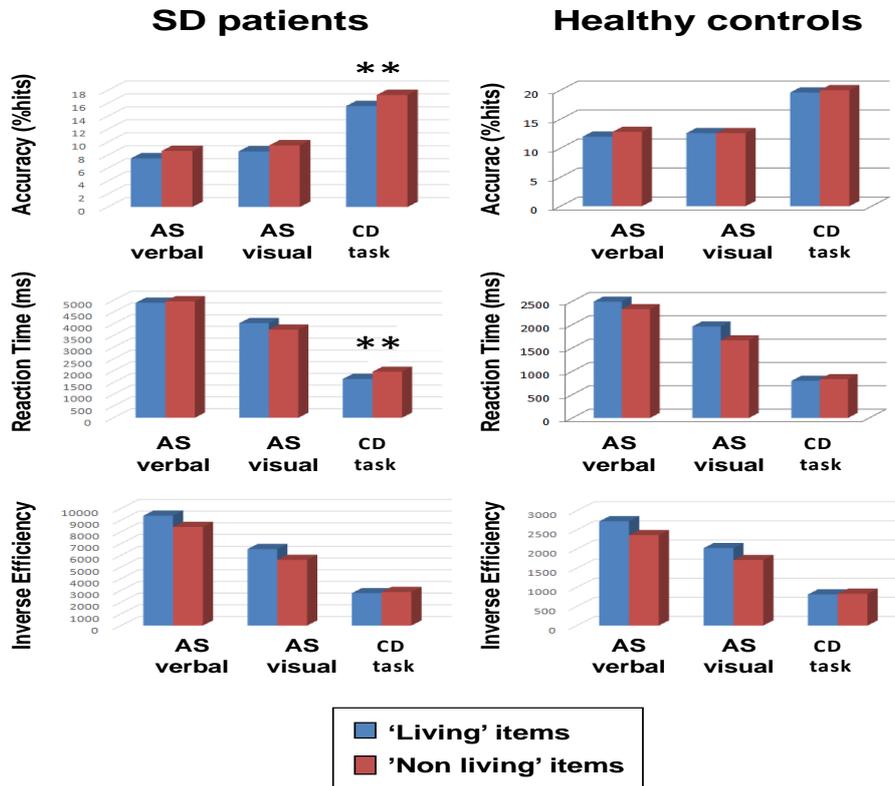


Figure 1 Histograms presenting behavioral outcomes (from top to bottom rows: Accuracy (% correct), Reaction Time for correct responses (in milliseconds, ms) and Inverse Efficiency for 'living' and 'non-living' items for the 3 tasks (AS verbal: AS visual: and CD). Significant differences ( $p < 0.05$ ) are indicated by \*\*.

## Discussion and conclusions

Our preliminary results failed to find robust evidence a differential processing of 'living' compared with 'non-living' category items in Sv-PPA patients. Correlational analysis between performances and regional FDG metabolism in temporal lobe yielded insightful but complex and overall statistically fragile outcomes. However, higher deficits for the categorization of "living" have been shown. The IE measurement yielded a cancellation of such significant differences, raising doubts on such category-specific dissociation in Sv-PPA reported for accuracy measures in the CD task. This outcome supports the view that a significant and reliable benefit for "non-living" reported elsewhere for Sv-PPA patients may fall under the 'exception' rather than be the rule (Lambon Ralph et al., 2007). Our results

allow to nuance the classical notion of the ATL as a trans-modal convergence ‘hub’ (Roger et al., 2017) and suggest that the two semantic categories could eventually be processed by different ATL sub-regions. The right superior temporal pole would act as a non-specific ATL region. The correlation with the parahippocampal region tied to performance in the AS visual task, which could be explained by the role of this area in contextual association retrieval (Baumann & Mattingley 2016). Moreover, assessing access to specific semantic categories with visual material only (words or pictures) in Sv-PPA patients affected by anomia and/or prosopagnosia might be prone to uncontrolled biases and potentially reflect difficulties in item identifying. Incorrect response does not unequivocally inform on a breakdown accessing its concept, but could reflect difficulties in identifying its structural attributes.

## References

- Baumann, O., Mattingley, J.B. 2016. Functional Organization of the Parahippocampal Cortex: Dissociable Roles for Context Representations and the Perception of Visual Scenes. *Linguistic Themes. The Journal of Neuroscience*, 36 (8), 2536–2542.
- Kondo, H., Saleem, K.S., Price, J.L. 2003. Differential connections of the temporal pole with the orbital and medial prefrontal networks in macaque monkeys. *Journal of Comparative Neurology*, 465,499-523.
- Lambon Ralph, M.A., Lowe, C., Roger, T.T. 2007. Neural basis of category-specific semantic deficits for living things: Evidence from semantic dementia HSVE and a neural network model. *Brain*, 130, (4), 1127-1137.
- Lambon Ralph, M.A., Jefferies, E., Patterson, K., Roger, T.T. 2017. The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience* 18, 42–55.

# Linguistic and extra-linguistic triggers in intrasentential codeswitches

Yuliya Leshchenko, Tatyana Ostapenko  
Perm State National Research University, Russia  
DOI:

## Abstract

The authors present an experimental study of bilingual intrasentential codeswitches considered in the context of the triggering hypothesis. Language shifts between two typologically distant languages with a low density of cognates – Komi-Permyak and Russian – are analyzed. The results prove that the switches are induced by several types of triggers. In the first instance, the study replicates the cognate facilitation effect (a word directly preceding or following a cognate increases the chances of a codeswitch). Secondly, it is shown that this effect depends on the direction of the codeswitch and is related to the dominance of the languages in bilingual consciousness. Finally, the influence of certain non-cognate triggers (the divergence of surface syntactic patterns along with the congruence of their underlying structures and the extra-linguistic/referential relatedness of the codeswitched words) is demonstrated.

Key words: bilingual word processing, codeswitching, triggering, cognate facilitation effect.

## Introduction

Current linguistics defines codeswitching as a concurrent use of two or more languages executed by a bilingual/multilingual speaker within one communicative context (Gardner-Chloros 2009). It is generally assumed that the degree of the speaker's control over language shifts can differ, so conscious and unconscious codeswitches are distinguished. Unconscious codeswitching is unintentional (not regulated by the speaker) and seems to be unmotivated. It commonly occurs in the intrasentential position: languages can be changed between clauses within a sentence or within clauses between constituents of a word phrase. The sources and reasons of unconscious codeswitching represent one of the most complicated and controversial issues in the general theory of bi/multilingual language processing that requires detailed investigation.

Though, at first sight, unconscious codeswitches seem to be occasional and fully unpredictable, there is a high possibility that they are governed by certain implicit psycholinguistic factors. According to M. Clyne's

hypothesis, unconscious shifts between languages can be triggered by certain words (trigger words) common for both of the speaker's languages – bilingual homophones or cognates with a high degree of formal and semantic overlap (Clyne 1980). Basing on the study of taped conversations of German-English and Dutch-English bilinguals, the author argues that “such words can cause speakers to lose their linguistic orientation and to codeswitch, either immediately after, or in anticipation” (ibid, p.401).

Further studies of the “cognate facilitation effect” (the presumption that the presence of a cognate greatly enhances the chances for a codeswitch) proved its validity for a pair of typologically related languages with a strong lexical overlap/high density of cognate words (Dutch and English) and two pairs of typologically distant languages with a weak lexical overlap (Moroccan Arabic and Dutch; Russian and English) (Broersma, de Bot 2006). Moreover, it has been hypothesized that, apart from cognate words, there may be other phenomena that can trigger a codeswitch. However, taken together, the general amount of data on triggering is scarce; much more studies of different language pairs are needed to elaborate the cognate facilitation hypothesis, to reveal some other triggering items and, overall, to prove that triggered codeswitching is a robust and general phenomenon.

### **Experimental research: Material and methodology**

In the paper, the authors present a study of unconscious intrasentential codeswitches performed by bilingual speakers of two typologically distant languages (Komi-Permyak and Russian) in experimental settings.

At the first stage of the study we conducted a psycholinguistic experiment (a free associative test) with adult Komi-Permyak – Russian bilingual speakers (65 university students aged 17 – 23). The test included two trials with the identical stimuli list of 54 high-frequency words presented first in the Komi-Permyak language, and then in Russian. As a result, we received a considerable number of “stimulus–reaction” pairs that make up codeswitching word combinations (CWCs) formed according to the syntactic rules of the Komi-Permyak and/or Russian languages (семья–ыджыт ‘large family’; чожа–быстро ‘walk quickly’). At the second stage, a survey revealing the usage frequency of the received CWCs in spontaneous speech of Komi-Permyak – Russian speakers was carried out.

### **Results and discussion**

Our study aimed at answering the following questions: Is cognate facilitation effect valid for the switches between the Komi-Permyak and Russian languages? Does this effect depend on the direction of the switch? What other factors (besides cognate words) can trigger a codeswitch?

To answer the first question, we analyzed how the presence of a cognate word in a CWC influences its overall usage frequency. The results of statistical analysis showed that the majority of the CWCs marked as habitual, regularly used by the Komi-Permyak–Russian speakers include a cognate word (either as the first, or the second constituent); the less frequent the CWCs are, the fewer of them contain a cognate. Thus, the cognate facilitation effect is obviously manifested in our experimental data.

Further, we examined how this effect depends on the direction of the codeswitch. For this purpose, we grouped all the high-frequency CWCs into the Komi-Permyak–Russian and Russian–Komi-Permyak ones and analyzed the position of the trigger within each CWC. Interestingly, we revealed that in Russian–Komi-Permyak CWCs the most frequent triggering type is linear triggering: the first constituent of a CWC is a cognate word (погода–бур ‘nice weather’). Alternatively, the majority of the Komi-Permyak–Russian switches are evoked by reverse triggering with the cognate word as the second constituent of a CWC (кывзыны–внимательно ‘listen attentively’).

Presumably, this tendency can be explained by the dominance level of the two languages in the speakers’ consciousness: though both are the native languages acquired from birth and used fairly often, Russian seems to dominate over Komi-Permyak for a number of reasons. Thus, it is highly probable that Russian words (words of the dominant language) have stronger intralingual connections, so that a switch to Komi-Permyak needs to be facilitated by a cognate word, as it requires certain effort on the part of the speaker. On the contrary, for Komi-Permyak words (words of the weaker language) both intra- and interlingual connections can be equally strong; therefore, switching to the Russian language does not need much cognate facilitation, as it is easy enough by itself.

Finally, the analysis of the CWCs with non-cognate constituents enabled to reveal two main factors that most probably evoked a codeswitch. In the first instance, this is the divergence of Russian and Komi-Permyak syntactic patterns on the surface structure level along with the congruence of their underlying structures. As long as Komi-Permyak has a bound word order which imposes certain syntactic constraints on word collocation (e.g., noun and verb phrases can be formed only with adjectives and adverbs in preposition), the speakers often conjoin a Komi-Permyak noun or verb with a Russian adjective or adverb as this does not contradict the Russian syntactic rules.

Another important factor that induced a codeswitch turned to be the extra-linguistic/referential one. The majority of non-cognate Komi-Permyak – Russian CWCs referred to the academic communicative context (вовлыны –на пары ‘attend lessons’, кывзыны–лекцию ‘listen to the lecture’), while most of the Russian–Komi-Permyak CWCs described family and friendly

relationship (любить–ай мамос ‘love Mum’; хороший–ёрт ‘good friend’). It should be noted that the triggering potential of the extra-linguistic factor has a dependent character: it can trigger a codeswitch by itself, when no other facilitating factors are present; however, when combined with the cognate facilitation effect/syntactic pattern divergence, its influence is boosted dramatically. Thus, the overwhelming majority (over 90%) of the highest-frequency CWCs were influenced by the cumulative impact of both linguistic and extra-linguistic triggers.

### **Conclusions**

The received results demonstrate that intrasentential codeswitches common for Komi-Permyak–Russian bilingual speakers can be induced by both linguistic and extra-linguistic triggers. The study replicated the existence of the cognate facilitation effect; besides, the relatedness of this effect to the dominance of the languages in bilingual consciousness has been shown. The influence of non-cognate linguistic (divergence of the surface syntactic patterns) and extra-linguistic (referential attributes of Komi-Permyak and Russian words) types of triggers on the codeswitch has also been demonstrated. Taken together, the received results contribute to a better understanding of the triggering mechanism in bilingual codeswitches.

### **Acknowledgements**

The research is supported by the Russian Foundation for Basic Research, grant №17-29-09074 “Combined trilingualism and its influence on linguistic and cognitive activity of an individual: an integrative model”.

### **References**

- Broersma, M., de Bot, K. 2006. Triggered codeswitching: A corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition* 9(1), 1–13.
- Clyne, M. 1980. Triggering and language processing. *Canadian Journal of Psychology* 34(4), 400–406.
- Gardner-Chloros, P. 2009. *Code-switching*. Cambridge, Cambridge University Press.

# Automatic detection of language impairment

Hali Lindsay<sup>1</sup>, Johannes Tröger, Nicklas Linz<sup>1</sup>, Jan Alexandersson<sup>1</sup>, Johannes Prudlo<sup>2</sup>

<sup>1</sup>German Research Center for Artificial Intelligence (DFKI), Germany

<sup>2</sup>Department of Neurology, University of Rostock and DZNE, Germany

DOI:

## Abstract

Amyotrophic Lateral Sclerosis (ALS) is a progressive neurodegenerative disease that presents with cognitive impairment (CI) in up to 50% of cases. In this study, we consider 17 healthy controls' and 61 ALS patients'—16 with CI and 45 without CI—performances on the cookie theft picture description task. Measures of language impairment are automated and used to evaluate the entire ALS population. Results show that both ALS patients with and without CI show statistically significant differences from healthy controls over four language features.

Key words: Amyotrophic Lateral Sclerosis, language impairment

## Introduction

Amyotrophic Lateral Sclerosis (ALS) is a muscle degenerative disease. An estimated 1/3 to 1/2 of cases present with mild to severe cognitive impairment (CI) primarily affecting executive functioning. Current literature suggests that mild language impairment is present—and may even be more prevalent than executive impairment—that cannot be attributed to declining motor function, social withdrawal, or executive dysfunction (Taylor et al., 2013). This paper aims to use automated, reproducible methods for identifying the presence of language impairment in ALS patients with and without CI.

## Methods

### Data

German Cookie Theft Picture descriptions from 61 ALS patients (age range=31–79 years; 22 female) and 17 age-, gender- and education-matched controls (age range=44–80 years; 6 female) were recorded and manually transcribed from audio files with Praat (Boersma & Weenink, 2019) by trained computational linguists following the CHAT protocol (MacWhinney, 1991). ALS patients were recruited from the university hospitals and DZNE sites in Rostock, Germany and Magdeburg, Germany and gave written consent; the study was conducted according to the Declaration of Helsinki and approved by the local ethics committees. Individuals with ALS were sorted into two groups; 16 with CI (*ALS<sub>ci</sub>*, age range=40–76 years; 7 female) and 45 without CI (*ALS<sub>noCI</sub>*, age range=31–79 years; 15 female). To screen for CI, raw

scores of the Montreal Cognitive Assessment (MoCA)—a common test used for screening cognitive ability—were compared to normalized German values for every subject. If a person scored below the fifth percentile they were considered for the *ALS<sub>ci</sub>* group (Strong et al., 2009). No healthy controls (HC) scored below the 5th percentile for the MoCA.

## Features

For general measures of production, the total number of words (*#<sub>words</sub>*) and the total number of sentences (*#<sub>sentences</sub>*) were counted from each sample. To evaluate language function, we considered four measures of performance:

(1) Mean words per sentence (*mean sentence length*) is calculated as a naive measure of sentence complexity.

(2) Mean dependency length (*mean dependency length*) is a proxy of syntactic complexity. The average dependency length over all utterances in a person’s CTP task is calculated. Dependencies are determined from the parse tree provided by the SpaCy German language models (Honnibal, Johnson 2015). A lower mean dependency compared to healthy controls would indicate difficulty in constructing complex sentences indicating language impairment.

(3) The ratio of open to closed class words (*open:closed*) is calculated where each word in a production is tagged with a part-of-speech (POS) using German SpaCy language models (Honnibal, Johnson 2015). Closed class words are considered to be conjunctions, demonstratives, prepositions, determiners, and pronouns; all other tags are considered as open class words. A high ratio for *open:closed* would be in line with agrammatic aphasia (Kim, Thompson, 2004) whereas a low ratio for *open:closed* may reflect the verb deficits identified by (Bak, Chandran 2011).

(4) Complexity index (*Complexity Index*) is a proxy of language ability that can be computed from spontaneous speech. Taylor et al. (2013) showed this measure to be useful in identifying language impairment in an ALS population. To calculate CI, responses are broken down into independent and dependent (subordinate or nonfinite) clauses (Hersen 2004). For this paper, the process was automated by using German language models from the Stanford parser (Rafferty, Manning 2008) to produce a parse tree for each response. Independent clauses were considered from the *root* tag and dependent clauses were collected from the *S* tag. The following was computed from each participant’s CTP task recording:

$$CI = \frac{S_{count} + Root_{count}}{Root_{count}}$$

## Results

For the statistical analysis, a one-way between-subjects Analysis of Variance was computed for each language feature for the HC,  $ALS_{no1}$  and  $ALS_{ci}$  groups. Significant effects across all groups were found for all language function features; *mean sentence length* ( $F(2,74) = 0.50, p < 0.001$ ); *mean dependency length* ( $F(2,74) = 0.38, p < 0.001$ ); *open:closed* ( $F(2,74) = 0.29, p < 0.05$ ); *Complexity Index* ( $F(2,74) = 0.44, p < 0.001$ ). No significant effects were found for  $\#words_i$  and  $\#sentences_i$ . Between-group posthoc analyses were conducted for all main effects with the Tukey HSD and yielded similar results. There were no significant differences between  $ALS_{no1}$  and  $ALS_{ci}$  but a significant difference between  $ALS_{no1}$  and HC and  $ALS_{ci}$  and HC. Significant results are visualised with box plots in Figure 1.

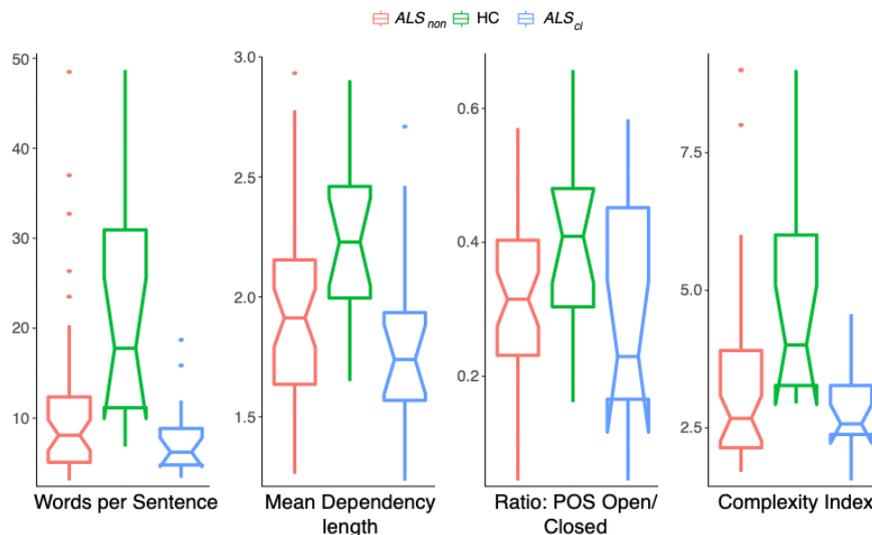


Figure 1. Box plots showing the group performance of each feature

## Discussion and conclusion

Overall, the  $ALS_{no1}$  group produces fewer complex responses than the HC group and performs on-par of those with  $ALS_{ci}$ . There is no significant main difference in  $\#words_i$  or  $\#sentences_i$  leading us to believe that this difference is primarily language and not due to a lack of overall production.

Taylor and colleagues (2013) stated that over 40% of patients impaired on language performed normally on tasks specific to executive impairment and suggested that these groups may even be dissociable. By comparing the  $ALS_{ci}$  and  $ALS_{no1}$  groups, there is evidence of language impairment in ALS that cannot be attributed to CI.

There is evidence, in line with current research—from automated analysis of language features—that both the *ALS<sub>no</sub>* and *ALS<sub>ci</sub>* groups show signs of language impairment. Future work should focus on the prognostic value of language impairment in terms of disease progression.

## Acknowledgements

Thank you to Mario Mina and Marc Sadler for assisting with analysis.

## References

- Bak, T., Hodges, J 2004. The effects of motor neurone disease on language: Further evidence. *Brain and Language* 89, 354-361.
- Boersma, Paul, Weenink, David 2019. Praat: doing phonetics by computer [Computer program]. Version 6.1, retrieved 13 July 2019 from <http://www.praat.org/>.
- Hersen, M. 2004. *Comprehensive Handbook of Psychological Assessment*. John Wiley & Sons, Inc.
- Honnibal, M., Johnson, M. 2015. An Improved Non-monotonic Transition System for Dependency Parsing.” Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP-2015), Lisbon, Portugal. D15-1162 1373-1378. doi: 10.18653/v1/D15-1162
- Kim, M., Thompson, C.K. 2004. Verb deficits in Alzheimer’s disease and agrammatism: implications for lexical organization. *Brain and language* vol. 88, 1 (2004): 1-20.
- MacWhinney, B. 1991. *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum Associates, Inc.
- Rafferty, A. Manning, Ch.D. 2008. Parsing Three German Tree. banks: Lexicalized and Unlexicalized Baselines. In *ACL Workshop on Parsing German*.
- Taylor, L., Brown, R., Tsermentseli, S., Al-Chalabi, A., Shaw, C., Ellis, C., Leigh, P., Goldstein, L. 2013. Is language impairment more common than executive dysfunction in Amyotrophic Lateral Sclerosis? *Journal of Neurology, Neurosurgery & Psychiatry*, 84, 494-498.
- Taylor, L., Brown, R., Tsermentseli, S., Al-Chalabi, A., Shaw, C., Ellis, C., Leigh, P., Goldstein, L. 2013. Is language impairment more common than executive dysfunction in Amyotrophic Lateral Sclerosis? *Journal of Neurology, Neurosurgery & Psychiatry*, 84, 494-498.
- Tröger, J., Linz, N., König, A., Robert, P., Alexandersson, J. 2018. Telephone-based Dementia Screening I: Automated SVF Assessment. Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth '18). ACM, NYC, NY, USA, 59-66.

# The use of pronominal forms in complement clauses in European and Brazilian Portuguese

Paula Luegi<sup>1</sup>, Marcus Maia<sup>2</sup>, Armanda Costa<sup>1</sup>

<sup>1</sup>School of Arts and Humanities, University of Lisbon, CLUL, Portugal

<sup>2</sup>Federal University of Rio de Janeiro, LAPEX, Brazil

DOI:

## Abstract

In this study, in a written sentence completion experiment, we contrast European and Brazilian Portuguese in what concerns the production of null, overt pronouns and full NPs, as subjects of verb complement clauses, to refer to the previously mentioned entities (or to extra-discursive referents). Results show that overall participants prefer to use null pronouns and that these pronouns are preferentially used to refer to the closest NP, especially in BP. Moreover, overt pronouns and full NPs are hardly used, being the later more frequent than the former, especially to refer to extra-discursive referents. Although the preference to refer to the closest subject with the null pronoun is in line with previous studies, high frequency of null pronouns and low frequency of overt ones in BP is not. We propose that these results might be explained by the characteristics of both the tested conditions and of the participants tested in our study.

Key words: European and Brazilian Portuguese, null and over subject pronouns, sentence completion task.

## Introduction

Although European and Brazilian Portuguese (EP and BP) share common properties, there are some differences between them, like the null subject parameter: Null subjects are not only less frequent in BP (Duarte 1995, Barbosa et al. 2005), compared to EP, but are also more restricted (Modesto 2018, Holmberg et al. 2009). This is especially true for 3<sup>rd</sup> person singular pronouns (Duarte & Figueiredo Silva 2016), the focus of the present study. Holmberg et al. (2009) propose that 3<sup>rd</sup> person null subjects in partial null subject languages, like BP, must have a local antecedent that c-commands it. This locality effect results in a rejection of co-reference between subjects that are not immediately adjacent, that is, when there is an intervening subject (Holmberg et al. 2009). Also, null subjects cannot refer to extra-discursive referents (that is, not explicitly mentioned). So, in (1) and (2), the antecedent must be *João* and cannot be *Pedro* or any other entity.

- (1) a. O João disse que [-] comprou um carro.  
b. O João disse que ele comprou um carro.  
“John said that (he)/he bought a car.”
- (2) O Pedro acha que o João disse que [-] comprou um carro.  
“Peter thinks that John said that (he) bought a car.”

Moreover, the interpretation of overt pronouns is also different in both varieties: While in EP null subject pronouns are preferentially interpreted as referring to the previous subject and the overt pronouns as referring to non-subject antecedents, in BP both null and overt subjects might be interpreted as referring to *João* both in (1a) and (1b).

However, all these descriptions are mainly based on intuitive judgments (see however Modesto (2018)) or in corpora analysis.

### **Experimental study**

In the present study we conducted an experiment with a written sentence completion task to test the descriptions presented in the previous section. Participants had to complete sentences like (3) and (4).

- (3) John said that \_\_\_\_\_.
- (4) Peter said that John knows that \_\_\_\_\_.

We aimed at: (i) compare the overall frequency of null and overt subjects in EP and in BP, (ii) verify the restrictions of null subjects reference, that is which antecedents are retrieved with null subjects: local, distant (intervening subject) or extra-discursive antecedents, and (iii) verify the differences between EP and BP in overt pronouns production and interpretation.

### **Participants, materials and procedure**

Twenty eight native speakers of EP, from the University of Lisbon, and 27 native speakers of BP, from the Federal University of Rio de Janeiro, participated voluntarily in this study after accepting the informed consent.

Experimental items, all main sentences with embedded complement clauses selected by the verb, were built in two different conditions: One-Referent condition (ex. (3)) and Two-Referents condition (ex. 4)). Sentences (24 experimental items and 48 fillers) were presented on OnExp one at a time in a randomized order in two lists in a Latin Square design. Participants had to complete the sentence after the word “that” and to answer the question: “Whom or what did you refer in the text you wrote?”.

### Analysis

Data was coded as to indicate (i) type of anaphoric expression produced: Null subject, Overt pronoun or full NP, and (ii) mentioned entity: Closest-NP (“John”, in (3) and (4)), Distant-NP (“Peter”, in (4)), or Extra-NP (Extra-discursive NP) if the referred entity was not in the previous clauses. Linear mixed models with binomial function and full maximal model were used considering type of anaphoric expression and mentioned entity as dependent variables in separated analysis.

### Results

When comparing Null subjects with Others (Overt pronouns and full NPs), there is a main effect of Variety (Table 2), with more Null pronouns in BP (Table 1), a main effect of Sentence Type, with more Null pronouns in One-Referent condition, and an interaction of Variety and Sentence Type, with more Null pronouns in One-Referent condition in BP. Contrasts of Null vs. Overt or vs. full NPs replicated these results. When comparing Overt with full NPs, there is a marginal effect of Sentence Type: less Overt pronouns in One-Referent condition (Est.=-7.972; SE=4.503; z=-1.771; p=0.076).

Table 1. Percentage of answers referring to Closest-NP, Distant-NP or Extra-NP with Null, Overt or full NP in One-Referent (EP: n=183; BP: n=250) or Two-Referents (EP: n=216; BP: n=226) conditions.

	One-Referent			Two-Referents		
	Null	Overt	full NP	Null	Overt	full NP
Closest-NP	64%	3%	0%	49%	9%	0%
EP Distant-NP	--	--	--	1%	8%	0%
Extra-NP	0%	1%	32%	0%	0%	32%
Closest-NP	82%	2%	0%	56%	6%	1%
BP Distant-NP	--	--	--	2%	13%	0%
Extra-NP	0%	0%	16%	0%	0%	22%

When comparing Closest-NP with all Others (Distant-NP and Extra-NP), there is a main effect of Variety (Table 2), with more references to the Closest-NP in BP, a main effect of Sentence Type, with more references to the Closest-NP in One-Referent condition, and an interaction of Variety and Sentence Type, with more references to the Closest-NP in One-Referent condition in BP.

Overall, participants tend to produce more null subjects than any other expression, especially in BP and in One-Referent condition. In both varieties, null subjects are mainly used to refer to the Closest-NP being the reference to Distant-NPs almost residual. Full-NPs are used especially to

refer to extra-discursive entities. There is a very low frequency of overt pronouns, in both conditions and in both EP and BP, being however more frequent in Two-Referents condition to refer to Distant-NPs.

Table 2. Results of the Linear Mixed Effects analysis.

	Anaphoric expression (Null vs. Others)				Retrieved antecedent (Closest vs. Others)			
	Est.	SE	z	p	Est.	SE	z	p
(Intercept)	0.84	0.28	2.99	0.003	1.08	0.28	3.87	<0.001
Variety	-0.90	0.42	-2.14	0.032	-0.81	0.42	-1.93	0.054
Sent. Type	1.11	0.24	4.71	<0.001	0.98	0.23	4.22	<0.001
Variety x Sent. Type	-0.80	0.40	-1.99	0.047	-1.03	0.41	-2.50	0.012

Null subject results are partially in line with previous proposals: Null subjects in BP are mainly used to refer to the Closest-NP (Holmberg et al. 2009, Modesto 2018). However, the high frequency of null subjects and the consequent low frequency of overt pronouns, in BP, is unexpected. This result might be explained by the fact that the conditions are structurally and discursively restricted leading to an increase of null pronouns production, especially in the 3<sup>rd</sup> person, more resistant to its replacement in BP (Barbosa et al. 2005). Also, or in addition, participants' characteristics might also explain this result: We tested university students, more likely to use the written norm and, therefore, to produce null subjects (see Almor et al. (2017) for similar results in written BP corpora). Further studies are necessary to explore not only the frequency of overt and null pronouns use but also the preference of interpretation of null subjects to retrieve Closest/Distant-NPs in comprehension tasks.

## References

- Almor, A., Maia, J.C., Lima, M.L.C., Vernice, M., Gelormini-Lezama, C. 2017. Language processing, acceptability, and statistical distribution: A study of null and overt subjects in Brazilian Portuguese. *Journal of Memory and Language* 92, 98-113.
- Barbosa, P., Duarte, E., Kato, M. 2005. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics* 4:2, 11-52.
- Duarte, E. 1995. A perda do princípio "evite pronome" no português brasileiro. PhD dissertation, Universidade Estadual de Campinas.
- Duarte, I., Figueiredo Silva, M.C. 2016. The null subject parameter and the structure of the sentence in European and Brazilian Portuguese. In W.L. Wetzels, J. Costa, S. Menuzzi (Eds.), *The Handbook of Portuguese Linguistics*, 234-253. Hoboken, NJ: John Wiley & Sons, Inc.
- Holmberg, A., Nayudu, A., Sheehan, M. 2009. Three partial null-subject languages. *Studia Linguistica*, 63(1), 59-97.

# Deep learning and intonation in Text to Speech systems

Philippe Martin

LLF, UFRL, Université de Paris, France

DOI:

## Abstract

Although Recurrent Neural Networks deliver excellent results in Speech-to-Text and Text-to-Speech (TTS) applications, the generation of satisfactory synthetic sentence prosody remains one of the main causes of the quality differences between human and synthetic speech. These differences do not involve only emotions or attitudes, but also the prosodic structure which determines the way the listener processes the speech flow. This paper explores the theoretical and technical reasons for these difficulties and proposes a better feature engineering approach for deep learning based on an alternate model of sentence intonation, applied to French.

Key words: Deep learning, artificial intelligence, text-to-speech, prosodic structure,

## TTS by unit selection

Most recent text-to-speech systems are based on the concatenation of segments of various sizes of actual human speech recordings (normally produced by one speaker) stored in a database. As complete required sentences are rarely found in the database, typical operations proceed by assembling selected segments and adjusting the syllabic prosodic parameters duration, fundamental frequency (pitch), and intensity. These parameters have to be predicted from the available information such as class of syllables and type of syntactic groups of words (syntagms), among others.

As dominant phonological models of sentence intonation (e.g. the Autosegmental-Metrical model), may or do appear confusing for speech engineers, many approaches rely on statistical prediction of segment duration and fundamental frequency patterns. These methods generally do not integrate any linguistic process such as the prosodic structure, but they incorporate constrained prosodic characteristics of human prosody linked to emotions, attitude and socio-geographic attributes.

## Deep learning: RNN, LSTM, BLSTM, DBLSTM

The advent of Recurrent Neural Network (RNN) and their variants such as Long Short-Term Memory (LSTM), Bidirectional BLSTM or Double bidirectional LSTM (DBLSTM) turn to be methods of choice used to generate appropriate sequences of prosodic parameters for TTS systems.

They can be viewed as generalizations of statistical approaches implemented earlier in TTS Systems. A deep learning implementation such as WaveNet (van den Oort et al. 2016) delivers synthesized speech of excellent quality very close to human speech. However, it generally fails to generate an appropriate prosodic structure from a given text, except for very short ones, where an acceptable pitch pattern is likely to be in the database.

Indeed, although the LSTM and BLSTM processes may capture time dependencies of some prosodic parameters, their practical implementation prevents them from apprehending long term dependencies characteristic of pitch movements from a raw wave signal. For instance, with a step size of 5 ms, 10 seconds of training speech would handle a memory of 2000 epochs, possibly involving a very large number of hidden layers. The number of connections to apply a gradient descent algorithm to determine connections between hidden layers could quickly become intractable. To cope with this problem, one has to turn to feature engineering, i.e. selection of features of the speech signal considered pertinent by prosodic experts.

### **Feature engineering**

One of the most important functions of the prosodic structure is to allow the listener to proceed to the decoding and the understanding of the message conveyed by the speaker. This may be difficult to operate by the listener in real-time from the text only. Indeed, the average short-term auditory memory is limited to some 2 to 3 seconds for speech. Therefore, it is imperative to process incoming speech quickly and efficiently. For this purpose, melodic contours located on stressed syllables (actually on stressed vowels) facilitate the process by segmenting the incoming speech flow and grouping the resulting chunks (the stress groups) hierarchically before the actual identification of the text. Melodic contours function as dependency markers to indicate to the listener how to assemble the successive stress groups in the sequence (which defines the prosodic structure).

The prosodic structures displayed on the synthesized and original example illustrated below were obtained automatically by the WinPitch software (2019) from the sequence of annotated melodic contours, following a dependency model. In this model, based on a generalization of Delattre (1966) *continuation mineure* and *continuation majeure*, but enacted with the contrast of melodic slope principle (Martin, 1975), the minor continuation contour  $C2 \searrow$  indicates a dependency towards a major continuation contour  $C1 \nearrow$  located further in the sentence (“to the right”), which in turn indicates a dependency towards a termination conclusive contour  $C0 \downarrow$ , also “to the right”. A neutralized contour  $C_n \rightarrow$  marks a dependency towards either  $C2 \searrow$ ,  $C1 \nearrow$  or  $C0 \downarrow$  located further in the sentence. Besides, a neutralized termination contour  $C0_n \leftarrow$  indicates a dependency relation this time “to the

left”, i.e. towards a terminal conclusive contour  $C0$  that precedes it. Details of the mechanisms defining the prosodic structure from a sequence of melodic contours can be found in Martin (2018).

Contrary to the original proposal by Delattre, melodic contours are not global, but are instead aligned on the vowel of (non-emphatic) stressed syllables. They are acoustically defined by their glissando values, above or below a threshold of change of pitch perception. Above this threshold, which depends on the rate of frequency change, the contour is perceived as a melodic change, below as a static tone (Rossi, 1971).  $C1\uparrow$  and  $C2\downarrow$  are respectively rising and falling contours above the glissando threshold, whereas  $Cn\rightarrow$  and  $C0n\leftarrow$  are below the threshold. The terminal conclusive contour  $C0\downarrow$ , falling in the declarative modality case and  $Ci\uparrow$  rising in the interrogative case, are acoustically defined as reaching respectively the lowest and the highest level of the pitch in the sentence.

### An example

A read speech sentence from the SIWIS corpus (FR\_A1\_08\_000 Yamagishi et al., 2017) has been selected as an example:

L'agriculture marocaine bénéficie d'un traitement privilégié pour ses exportations vers l'Europe “Moroccan agriculture benefits from privileged treatment for its exports to Europe”.

API segmentation, melodic contours assignment and prosodic structure generation are automatically generated by WinPitch. The original realization read by a female speaker is shown Fig 1, whereas Fig. 2 displays the realization of a TTS system (Claude voice in Microsoft Windows).

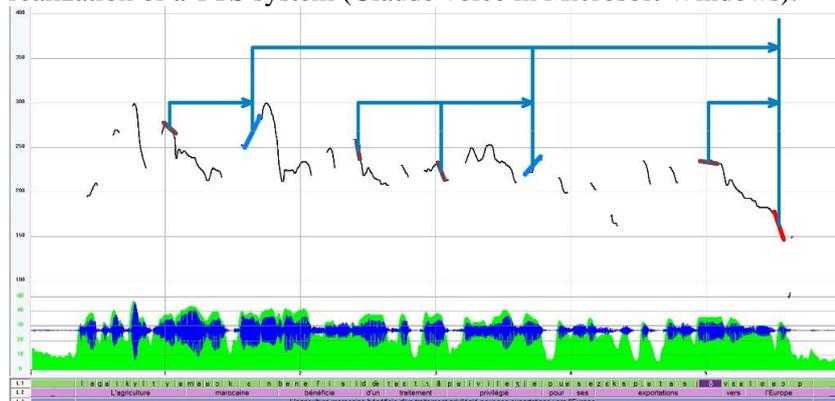


Figure 1. Original version of the sentence *L'agriculture marocaine bénéficie d'un traitement privilégié pour ses exportations vers l'Europe* read by a female speaker. The order of processing by the listener, as indicated by the prosodic structure, is [*L'agriculture*  $Cn\rightarrow$  *marocaine*  $C1\uparrow$ ], [*bénéficie*  $Cn\rightarrow$

*d'un traitement Cn→ privilégié C1↗*] and [*pour ses exportations Cn→ vers l'Europe C0↓*].

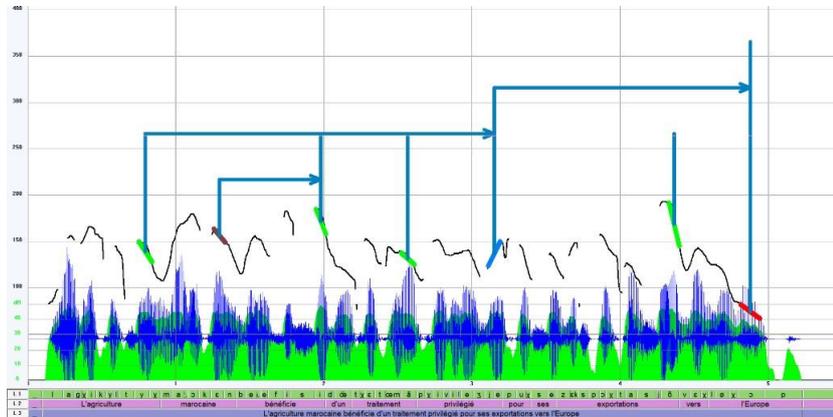


Figure 2. TTS version of the sentence *L'agriculture marocaine bénéficie d'un traitement privilégié pour ses exportations vers l'Europe* (Microsoft voice Claude). For this TTS realization, the order of processing by the listener, as indicated by the prosodic structure, is [*L'agriculture C2↘ [marocaine Cn→ bénéficie C2↘] d'un traitement C2↘ privilégié C1↗*] and [*pour ses exportations C2↘ vers l'Europe C0↓*]. The falling penultimate melodic contour C2↘ is a-grammatical, and should have been realized as a neutralized contour Cn→, below the glissando threshold.

## References

- Delattre, P. 1966. Les dix intonations de base du français, *French Review* 40, 1-14.
- Martin, Ph. 1975. Analyse phonologique de la phrase française, *Linguistics*, 146, 35-68.
- Martin, Ph. 2018. Intonation, structure prosodique et ondes cérébrales, London, ISTE, 322 p.
- Rossi M. 1971. Le seuil de glissando ou seuil de perception des variations tonales pour la parole, *Phonetica* 23, 1-33.
- van den Oord, A. et al. 2016. A Generative Model for Raw Audio, <https://arxiv.org/abs/1609.03499>
- Yamagishi, J., Honnet, P-E., Garner, Ph., Lazaridis, A. 2017. The SIWIS French Speech Synthesis Database, 2016 [dataset]. University of Edinburgh. School of Informatics. The Centre for Speech Technology Research. <https://doi.org/10.7488/ds/1705>.
- WinPitch. 2019. [www.winpitch.com](http://www.winpitch.com)

# Transcranial stimulation in semantic dementia

Elsa Mhanna<sup>1</sup>, Clara Sanches<sup>1</sup>, Antoni Valero-Cabré<sup>1</sup>, Marc Teichmann<sup>1,2</sup>

<sup>1</sup>Institut du Cerveau et la Moelle Épineière, Sorbonne Université, France.

<sup>2</sup>Department of Neurology, Assistance Publique Hôpitaux de Paris, France.

DOI:

## Abstract

Semantic variant of primary progressive aphasia (sv-PPA) is the most frequent form of PPA to with no current treatment. Non-invasive brain stimulation by transcranial Direct Current Stimulation (tDCS) is emerging as a therapeutic alternative. We present preliminary results from a double-blind sham-controlled study of 14 sv-PPA patients who received daily tDCS sessions for 10 consecutive days. Patients were randomized to 3 conditions: Left Anodal (excitatory), Right Cathodal (inhibitory) and Sham (Placebo). Participants underwent series of language tasks, prior and following the sessions. Prior and 2 weeks following the tDCS regime patients underwent <sup>18</sup>[F]-FDG-PET recordings. We didn't find significant improvements in semantic access for any group. Left Anodal group showed medium effect sizes in visual semantic associations. Excellent tolerance and a high level of subjective satisfaction was found for all groups. The protocol is ongoing.

Key words: Semantic variant, primary progressive aphasia, repetitive tDCS

## Introduction

The semantic variant of primary progressive aphasia (sv-PPA) or Semantic Dementia is the most frequent of all PPA forms and its main feature is a loss of conceptual knowledge with preserved fluency and repetition (Gorno-Tempini 2004). Little can currently be done to effectively treat sv-PPA patients.

Transcranial direct current stimulation (tDCS) has been employed in neurodegenerative diseases with encouraging results (Cotelli 2014). This technology delivers a weak (1-2 mA) constant current between two electrodes (an anode and a cathode) placed on separate scalp locations.

Results from a pre-therapeutic study (Teichmann 2016) evaluated the impact of a single tDCS session on a cohort of 12 sv-PPA patients. The current protocol is a prospective double-blind randomized clinical trial evaluating the impact of a periodical regime of daily tDCS sessions. We compared the impact of 2 stimulation strategies: Anodal (excitatory) tDCS on the left ATL (Anterior temporal lobe) to boost impaired systems of the left hemisphere involved in semantic access and, Cathodal (inhibitory) tDCS, targeting the right ATL, to suppress interhemispheric inhibitory interactions and a Sham (placebo).

## Material and methods

14 patients diagnosed with sv-PPA were enrolled in our study at the Pitié-Salpêtrière University Hospital. Diagnosis of sv-PPA was made according to international criteria (Gorno-Tempini et al., 2011).

A cohort of healthy controls (HC, n=5) were matched with our patients for gender, age, socio-cultural level and handedness (differences for all matched variables between sv-PPA and HC  $p > 0.05$ ). Patients were randomized into 3 groups. All gave informed signed consent.

Table 1: Demographical information about the participants in the study.

Tests	LA tDCS	RC tDCS	Sham tDCS	HC
Nr of Participants	5	4	5	5
Sex M/F	3/2	3/1	3/2	4/1
Age (years)	68,6 ( $\pm 5,3$ )	66,7 ( $\pm 5,1$ )	73,4 ( $\pm 2,1$ )	64,6 ( $\pm 6,9$ )
Disease duration (years)	5 ( $\pm 1,4$ )	3,7 ( $\pm 0,7$ )	4 ( $\pm 1,7$ )	-

LA: Left anodal; RC: Right Cathodal; HC: Healthy Controls.

This clinical trial is a double-blind sham-controlled study involving 10 daily tDCS sessions over 2 weeks. All behavioral evaluations were performed at 4 time-points of the clinical follow up: pre tDCS baseline, 3 days after the 10<sup>th</sup> and final tDCS session, 2 weeks later and 4 months later.. Healthy controls were only evaluated once following identical behavioral tasks and neuroimaging protocols as sv-PPA patients. To deliver tDCS we used a device by 'NE STARSTIM' model 'ENOBIO' 8 channels. During active anodal or cathodal stimulation (LA and RC groups), current was leveled up to reach an intensity 1.59 mA (current density: 0.06 mA/cm<sup>2</sup>) (Teichmann 2016). Stimulation lasted 20 minutes.

A verbal and visual semantic association tasks that include 78 words (SA verbal) or pictures (SA visual) grouped in 26 trials each assembling 3 words or pictures in which only 2 are related semantically.

All patients underwent <sup>18</sup>[F]-FDG-PET (<sup>18</sup>[F]-Fluoro-2-Desoxy-2-Fluoro-D-Glucose Positron Emission Tomography) at baseline and 2 weeks after the end of the 10 days tDCS treatment to document and quantify lasting modifications of temporal lobe metabolism by the stimulation regime.

## Results

The three groups of sv-PPA patients achieved significantly lower accuracy levels than their matched healthy controls (HC) at inclusion. A non-parametric Wilcoxon test applied to each comparison yielded statistically

significant results ( $p < 0.05$ ) for all three comparisons (LA vs HC, RC vs HC and sham vs HC) (See Figure 1).

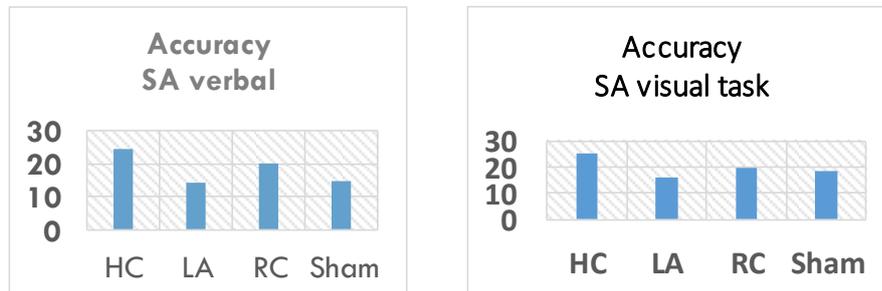


Figure 1. Mean accuracy for the verbal Semantic Association (SA verbal) and visual Semantic Association (SA visual) tasks measured at inclusion in healthy controls (HC,  $n=5$ ) and the sv-PPA patients randomized to Left Anodal tDCS (LA,  $n=5$ ), Right Cathodal tDCS (RC,  $n=4$ ) or Sham tDCS (Sham,  $n=5$ ) groups.

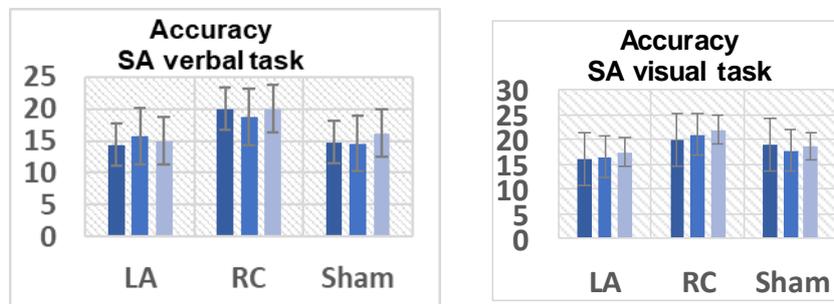


Figure 2. Preliminary outcomes calculated for three time-points across our follow up for Semantic association tasks. Data are presented for Baseline, 3 days post-stimulation (Post 1) and 2 weeks after tDCS offset (Post 2) and separated for each tDCS group: Left Anodal tDCS (LA), Right Cathodal tDCS (RC) and Sham tDCS groups. Time points for each tDCS group (LA, RC and SHAM) are presented serially in blue hues (from dark to light blue, Baseline, Post 1 and Post 2) in correlative order.

We performed non-parametric paired and unpaired tests to identify potential effects of tDCS stimulation based on the tasks. We compared post vs. pre tDCS regime accuracy shifts (Wilcoxon rank test) for each individual stimulation groups. We also explored significant differences between *post minus pre* accuracy differences, comparing active tDCS (LA, RC) with sham tDCS (SHAM) groups (U-Mann Withney test). It yielded no significant

results for any tDCS groups (All Wilcoxon rank tests  $p > 0.05$ ). Similarly, comparison of post *minus* pre tDCS accuracy differences across groups did not reveal any statistically significant difference (All U-Mann-Withney tests  $p > 0.05$ ). See Figure 2. Potentially interesting, a Cohen's *d* test showed a mild effect size for Left Anodal tDCS stimulation ( $d = 0.46$ ) in the SA visual task compared with the other tDCS groups (RC and SHAM).

Finally, we attempted to identify a metabolic  $^{18}\text{F}$ -FDG-PET biomarker at inclusion able to predict improvements driven by active tDCS treatment. One significant correlation was noted between metabolism in the right middle temporal pole and SA visual task improvements in both active tDCS groups ( $p$  value of 0.04,  $r = 0.69$  Pearson correlation).

## Discussion and conclusion

Thus far, we can conclude that our 10 days tDCS regime showed to be well tolerated and no stimulation negative side-effects were noted in concordance with our pilot study (Teichmann 2016). Even if very preliminary, correlation between the semantic tasks and the  $^{18}\text{F}$ -FDG-PET metabolic levels showed a statistically significant correlation between FDG metabolism in the right middle temporal pole and post-pre tDCS outcomes for the SA visual task. Such observation could suggest a potential role for the right hemisphere subtending in visual semantic processes. On the other hand, we were unable yet to find clear evidence supporting the beneficial effects of active tDCS stimulation on the language tasks. However, the medium effect size found for the left anodal tDCS group (LA) compared to the other two (RC and Sham) suggests that with a higher number of patients statistically significant differences might arise between active tDCS and sham groups.

## References

- Gorno-Tempini M.L., Dronkers N.F., Rankin K.P., Ogar, J.M., Phengramy, L., Rosen, H.J. Johnson, J.K. Weiner, M.W., Miller, B.L. 2004. Cognition and anatomy in three variants of primary progressive aphasia. *Ann Neurol* 2004, 55, 335-346.
- Cotelli, M., Manenti, R., Petesi, M., Brambilla, M., Cosseddu, M. Zanett, O. Miniussi, C., Padovani, A., Borroni, B. 2014. Treatment of Primary Progressive Aphasias by Transcranial Direct Current Stimulation Combined with Language Training. *Journal of Alzheimer's Disease* 39, 799-808.
- Teichmann M., Lesoil C., Godard J., Vernet, M., Bertrand, A., Levy, R. Dubois, B., Lemoine, L., Truong D.Q., Bikson, M., Kas, A., Valero-Cabre, A. 2016. Direct current stimulation of temporal poles boosts semantic 93-707. *Processing in primary progressive aphasia. Annals of Neurology* 2016, 80 (5): 6.
- Valero-Cabr e, A., Sanches, C., Godard, J., Fracchia, O., Dubois, B., Levy, R., Truong, D.Q., Bikson, M., Teichmann, M. 2019. Language boosting by transcranial stimulation in Progressive Supranuclear Palsy. *Neurology* 2019.

# **Semantic abilities and narration in aphasia in Greek**

Fani Nasika

Hellenic Open University, Greece

DOI:

## **Abstract**

This paper aims to present the findings of clinical linguistic analysis on narrative speech in aphasia. Aphasia is an acquired language disorder caused by brain lesions and it can be distinguished into non-fluent and fluent aphasia. Both categories show linguistic deficits affecting language production and/or comprehension. Aphasic narrative speech is considered particularly impaired in aphasia in respect of morphosyntax and semantics (Goodglass, Kaplan, Barresi 2001). Thus, in this study, the semantic abilities of four aphasic Greek speakers, two with non-fluent and two with fluent aphasia, were examined during storytelling. The samples were analyzed semantically with the use of PRISM (Crystal 1992). The results showed the semantic abilities of the participants both at a lexical level and at thematic structure level.

Key words: PRISM, semantic analysis, aphasia, Greek.

## **Theoretical background**

Aphasia is an acquired language disorder which can affect skills related to language expression and/or comprehension, thus highly compromising communication (Goodglass, Kaplan, Barresi 2001), while it is caused by lesions in the brain, following a cardio-vascular accident (CVA), a tumour, a traumatic brain injury (TBI) or other neurological causes. These lesions should be located in the language-dominant hemisphere. Moreover, aphasia can be distinguished into two main aphasia types: the non-fluent and the fluent aphasia. This distinction has been based on the fact that the most prominent feature differentiating these two types of aphasia is fluency (Goodglass, Kaplan, Barresi 2001).

With reference to non-fluent aphasia, following Goodglass, Kaplan and Barresi (2001), it is characterized by non-fluent speech production, reduced or ungrammatical sequences, pauses, and difficulties with articulation, inflection and word retrieval. In relation to fluent aphasia, individuals have word finding difficulties, especially with nouns, and their comprehension is impaired. Their errors include omissions and substitutions as well as paraphasias. Finally, they may present with difficulties in thematic role assignment both in production and comprehension (Caramazza, Miceli 1991).

In this respect, it seems that both individuals with non-fluent and fluent aphasia face difficulties with narrative speech, although these difficulties may be of different nature (Goodglass, Kaplan, Baresi 2001), thus expecting more morphosyntactic errors in non-fluent aphasia and more semantic errors in fluent aphasia. A tool for profiling these abilities is PRISM (Crystal 1992). PRISM comprises two separate procedures: PRISM-L, which examines the relationship between semantics and the mental lexicon, and PRISM-G, which examines the relationship between semantics and grammar (i.e., thematic-role structures). In the present study, PRISM-L and PRISM-G were used to profile the semantic abilities of Greek-speaking individuals with aphasia during the narration of a story.

## **Methodology**

### **Sample**

The narrative speech samples were collected from four participants, two with non-fluent aphasia and two with fluent aphasia ( $M = 56.5$ ;  $SD = 19.8$ ;  $range = 39-77$ ). There were three male and one female participant. All participants had suffered a CVA at least 6 months before participating in this study. In order for them to participate, they were assessed with the Greek standardized adaptation of the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass, Kaplan 1972; for the Greek adaptation: Papathanasiou, Feidantsi, Katsantoni, Panagiotopoulou, Malefaki 2004).

### **Material and procedure**

The speech samples were collected through story-telling, with the use of a Cinderella story booklet with pictures (Berndt, Wayland, Rochon, Saffran, Schwartz 2000). The collected samples were analyzed with the use PRISM (Crystal 1992), which was adapted into Greek (Nasika, Afantenou, Kremmyda 2017). More specifically, the samples were analyzed at a lexical level with PRISM-L and at a thematic structure level with PRISM-G.

### **Results**

Regarding participant 01, he was non-fluent and, in his narration, he used words from 24 out of 61 semantic fields, namely man, clothing, moving, making/doing, sight, language, quantity, time, etc. Of these semantic fields, the ones with the smallest lexical range were making/doing, happening, sight, and place. Concerning PRISM-G, he produced mainly clauses of two (62%) or three (22%) semantic elements, which places him in Stages II and III. The most frequent combination in Stage II was that of a dynamic verb with a theme element (65%) while in Stage III the most frequent combination was that of a dynamic verb with a theme and a temporal,

locative or other element (55%). Very often, he omitted the Actor of the action or the Experiencer.

Referring to participant 02, he was non-fluent and he produced words from 20 out of 61 semantic fields while not producing words from the last 12 semantic fields of PRISM-L, which correspond to later stages of lexical development. The semantic fields covered were mostly man, body, clothing, moving, tools, animals, food, make/do, etc. Of these semantic fields, the ones with the smallest lexical range were body and clothing. Regarding PRISM-G, participant 02 produced clauses with two (71%) or three (29%) semantic elements, which places him in Stages II and III. The main thematic structure used was that of a dynamic verb with a theme element. The Actor element was used only once.

As for participant 03, she had fluent aphasia. She used words from 14 out of 61 semantic fields, namely man, clothing, moving, making/doing, happening, place, time, etc. Of these semantic fields, the one with the smallest lexical range was body. Concerning PRISM-G, she produced mainly clauses of one (32%) or two (58%) semantic elements, which places her in Stages I and II. The most frequent structure in Stage I was that of a dynamic verb (83%) while in Stage II the most frequent combination was that of a dynamic verb with a theme (45%). Very often, she omitted the Actor of the action.

Concerning participant 04, he also had fluent aphasia. He produced words from 22 out of 61 semantic fields, namely man, clothing, moving, making/doing, happening, feeling, quantity, place, time, etc. Of these semantic fields, the ones with the smallest lexical range were man and clothing. Concerning PRISM-G, he produced mainly clauses of two (70%) or three (18%) semantic elements, which places him in Stages II and III. The most frequent structure in Stage II was that of a dynamic verb with a theme (30%) while in Stage III the most frequent combination was that of an actor + dynamic verb + theme (33%).

## Conclusions

The profiling of the semantic abilities of the four participants with the use of PRISM showed similarities irrespective of the type of aphasia. More specifically, regarding PRISM-L, it was shown that all participants used on average 33% of the available semantic fields while there was variation concerning the lexical range of the fields used. In relation to PRISM-G, most participants used mainly two and three-element thematic structures, corresponding to Stages II and III, while one participant with fluent aphasia used mainly structures from Stage I and Stage II. The thematic structure mostly used was that of a dynamic verb + theme. It should be noted that the Actor element was frequently omitted. However, this omission does not lead

to an ungrammatical structure in Greek but it may have consequences for the cohesion and the coherence of the narration.

Overall, it can be seen that the participants used a small range of semantic fields while they used simple thematic structures which correspond to early developmental stages. Thus, it seems that their semantic abilities are compromised and they bear certain similarities which should be further studied. In addition, it can be seen that PRISM is a useful tool for profiling the semantic abilities of individuals with aphasia, thus, providing information for clinical assessment and intervention.

## References

- Berndt, R.S., Wayland, S., Rochon, E., Saffran, E., Schwartz, M. 2000. Quantitative production analysis. Training manual. Hove, Psychology Press.
- Caramazza, A., Miceli, G. 1991. Selective Impairment of Thematic Role Assignment in Sentence Processing. *Brain and Language*, 41, 402-436.
- Crystal, D. 1992. Profiling linguistic disability. London, Whurr Publishers.
- Goodglass, H., Kaplan, E. 1972. *The Assessment of Aphasia and Related Disorders*. London, Henry Kimpton Publishers.
- Goodglass, H., Kaplan, E., Barresi, B. 2001. *The Assessment of Aphasia and Related Disorders*. 3<sup>rd</sup> Edition. Lippincott Williams & Wilkins.
- Nasika, F., Afantenou, K., Kremmyda, G. 2017. Using PRISM to profile semantic ability in Greek. In Botinis, A. (ed.), *Proc. 8<sup>th</sup> Tutorial & Research Workshop on Experimental Linguistics*, 77-80, Heraklion, Crete, Greece.
- Papathanasiou, I., Feidatsi, M., Katsantoni, M., Panagiotopoulou, E., Malefaki, S. 2004. The validation of Boston Diagnostic Aphasia Examination (BDAE-3) in Greek. Paper presented at the 26th World Congress of the International Association of Logopaedics and Phonetics, 29 August-2 September 2004, Brisbane Australia.

# Narrative development and gestures in different genres

Gökhan Özkayin

Institute of German Studies, University Koblenz-Landau, Germany

DOI:

## Abstract

This paper explores genre-specific diversities of verbal narratives in different situational contexts. The main purpose is to analyse the narrative development and the usage of manual gestures of German speaking children in the age of 6-10. Data analysis revealed that the quantity of co-speech gestures differs not only with the age of the speaker and the situational contexts but also with narrative genres.

Key words: narrative development, gestures, children, spoken language, multimodality

## Introduction

The narrative genre or the inducement to tell a story causes and influences not only the structure but also recurrent patterns in children's narrations. Studies on gestures in discourse (e.g., McNeill 2005; Gullberg 2006) have shown that speakers build coherent discourse by using speech and co-occurring gestures during narratives.

The focus of this research is to analyse whether and to what extend different narrative genres effect the usage of gestures.

## Method

The presented results are related to one of several experiments of an ongoing study. All experiments are recorded on video. The examined videos of children's narratives are recorded in two different settings: 1. presentations of self-created stories told as monologues in front of their classmates. 2. interactional narratives told to a single classmate. In the first setting the narrators tell picture stories or story continuations. In the second setting the variants of the stimuli are divergent. Children retell either an animated short movie with no verbal stimuli or they tell the continuation of a story they just listen to, so without visual stimuli.

Following stimuli were used in the depicted experiment below: 1. an animated short movie without speech. 2. An instruction to tell self-experienced narratives or to tell the continuation of a story they hear.

This leads to following contrastive situational contexts that are analysed in this paper: (i) monologues vs. interactional narratives (ii) fantasy stories

vs. self-experienced narratives. A wider scope of this study is on the one hand the investigation of verbal cohesive elements of speech that create a coherent narration and on the other hand the non-verbal or implicational (e.g., bridging, presuppositions or framing) creations of coherent units. Several studies characterize these elements as one of the main indications for narrative development of children in elementary school (e.g., Boueke et al. 1995).

The variations in the data lead to investigations of several, related aspects of the interplay between gestures and verbal language, such as the syntax of discourse referents and possible co-referents as well as the effectiveness of cohesive gestures to create a coherent narration.

### **Participants**

The investigations are based on a video-corpus of German-speaking children from elementary school that tell or retell different types of narrations. The videos were recorded and collected at one elementary school in Germany (Rheinland-Pfalz) in 2019. The 68 participating children were in the age of 6-10. Distribution of children by age: 6-7 (15 participants); 7-8 (12 participants); 8-9 (20 participants) and 9-10 (21 participants).

### **Gestures**

A huge number of approaches exist to describe or define a gesture. In this study gesture is defined as the visible movement of any body part consciously or unconsciously made with the intention of communicating while speech is being produced (Calibris 2011). Still the focus in this experiment described below is on manual gestures. The combined use of speech and gesture develops, for example in the increase of rhythmic beat gestures with increasing narrative sophistication (Colletta 2004).

That's why special emphasis is on the quantity of discourse structuring gestures and co-speech iconic gestures. Different terminologies exist to describe the phenomenon of discourse structuring gestures: e.g., *beats* (McNeill 1992), *discourse gestures* (Müller 1998) or *gestures with a parsing function* (Kendon 2004). The term beat is preferred in this study. A simplistic definition of beat gestures describes them as non-referential and displaying a rhythmic behavior that is often accompanied by prosodic phenomenon. They also can assume a wide range of pragmatic functions. Whereas the term iconic gesture (McNeill 1985) refers to hand gestures that represent meaning that is related to the semantic content of the segments of speech they accompany.

### Data

Because of significant differences in the length of the narratives the presented numbers display the average quantity of manual gestures per 50 words in a narrative.

Table 1. Average quantity of manual beat gestures per 50 words in retellings of an animated short movie

Age	Monologue	Interactional
6-7	0.21	0.34
7-8	0.28	0.48
8-9	0.44	0.89
9-10	0.71	1.02

Table 1 shows the increase of manual beat gestures with the age of the speaker. Apart from this the quantity of beats also increase in interactional situations when they tell a story to a single classmate who is indicating a prompt response during the interaction.

Table 2. Average quantity of iconic gestures per 50 words in retellings of fantasy stories vs. self-experienced narratives

Age	Fantasy	Self-exp.
6-7	1.12	1.45
7-8	1.37	1.47
8-9	1.89	2.23
9-10	1.92	2.54

Table 2 shows the increasing of iconic gestures with the age of the speaker in fantasy stories as well as in self-experienced stories. Even if the amount of iconic gestures is significantly higher in self-experienced stories. Table 1 and 2 also shows that children were able to adapt their narration to the addressee and to the communicative context.

### Results

It can be seen from the collected data that the narratives and especially the quantity of co-occurring gestures vary not only in different situational contexts but also with the application of diverse stimuli. Furthermore, the results of the study demonstrate that narrative development is accompanied by gestural development, too. Older children use manual gestures more often

than younger children to structure information, for attributional uses, to introduce new discourse-referents or to refer to available discourse-referents.

The analyses demonstrate effects of cohesive gestures on subsequent discourse comprehension and production in children's narratives as it has been shown on other languages (for English: Goodrich Smith/Hudson Kam 2012; for Japanese: Sekine/Kita 2015).

Children's use of gestures related to narrative organization becomes more developed as children's narratives become more complex (Colletta et al. 2015). Additionally, gestures help children to boost their subsequent narrative performance.

## References

- Boueke, D., Schüle, F.; Büscher, H. 1995. *Wie Kinder erzählen. Untersuchungen zur Erzähltheorie und zur Entwicklung narrativer Fähigkeiten*. München, Fink.
- Calbris, G. 2005. *Elements of meaning in gesture*. Amsterdam, John Benjamins.
- Colletta, J. M. (2004). *Le développement de la parole chez l'enfant âgé de 6 à 11 ans*. Corps, langage et cognition. Sprimont, Mardaga.
- Colletta, J.-M., Guidetti, M., Caprici, O., Cristilli, C. 2015. Effects of age and language on co-speech gesture production: an investigation of French, American, and Italian children's narratives. *Journal of child language* 42, 122–145.
- Goodrich Smith, W., Hudson Kam, C.L. 2012. Knowing 'who she is' based on 'where she is': The effect of co-speech gesture on pronoun comprehension. *Language and Cognition*, 4-2, 75-98.
- Gullberg, M. 2006. Handling Discourse: Gestures, Reference Tracking and Communication Strategies in Early L2. *Language Learning*, 56, 155-196.
- Kendon, A. 2004. *Gesture. Visible Action as Utterance*. Cambridge, University Press.
- McNeill, D. 1985. So you think gestures are nonverbal? *Psychological Review* 92 (3), 350–371.
- McNeill, D. 1992. *Hand and Mind. What Gestures Reveal about Thought*. Chicago, The University of Chicago Press.
- McNeill, D. 2005. *Gesture and Thought*. Chicago, University of Chicago Press.
- Müller, C. 1998. *Redebegleitende Gesten. Kulturgeschichte – Theorie – Sprachvergleich*. Berlin, Arone Spitz GmbH.
- Sekine, K., Kita, S. 2015. Development of multimodal discourse comprehension: cohesive use of space by gestures. *Language, Cognition and Neuroscience*, 30, 1245-1258.

# Simulating alphabet recitation under thalamic lesions

Martin D. Pham<sup>1</sup>, Terrence C. Stewart<sup>2</sup>, Randy Harris<sup>3</sup>

<sup>1</sup>The Centre for Computational Medicine, The Hospital for Sick Children, Canada

<sup>2</sup>Applied Brain Research, Canada

<sup>3</sup>Department of English Language and Literature, University of Waterloo, Canada

DOI:

## Abstract

We utilize the Semantic Pointer Architecture, a neurocognitive architecture in order to model language impairments. Constructed is a spiking neural network to investigate the effect of neural deficits in the basal ganglia and thalamus on the retrieval of an ordered sequence of unique symbols. The model includes four subnetworks: associative memory, working memory, basal ganglia and thalamus. A lesion is simulated by reducing the number of available neurons in the thalamus and attenuating its input from the basal ganglia. The model remains mostly successful in the ordered retrieval of the alphabet but ‘stutters’: working memory ‘forgets’ the current letter and ‘steps back’ several letters before continuing correctly.

Key words: spiking neural network, language, lesions

## Introduction

The model makes use of the Neural Engineering Framework (NEF) (Eliasmith, 2004) methodology and Semantic Pointer Architecture (SPA) (Eliasmith, 2013) to construct a spiking neural network in the neural simulator Nengo (Stewart, 2009). NEF is an approach to neural simulation with an emphasis on three principles: neural representation, transformation and dynamics; each principle related to a mathematical formulation. SPA is an approach to cognitive modelling that represents higher-level cognitive concepts within the NEF as high-dimensional vectors.

## Neural Engineering Framework

The NEF is a framework for understanding neurobiological systems within a rigorous mathematical formulation consisting of three principles relating the former to the latter:

### Representation

Neuronal activations in a population of neurons encode a time-dependent vector signal nonlinearly as a spike train and decode as a weighted linear combination.

**Transformation**

Decoding of an input signal can estimate some nonlinear transformations.

**Dynamics**

Neural representations can be considered as control theoretic state variables. Neural dynamics can thus be described and analyzed mathematically using control theory.

The NEF may be used to design spiking neural networks to model perceptual, motor and cognitive systems. The software Nengo implements these principles to construct biologically plausible neural systems.

**Semantic pointer architecture**

The SPA is a framework that treats higher-level cognitive functions using neural representations carrying partial semantic content that may be composed into complex representational structures using basic mathematical operations. By representing concepts as high-dimensional vectors within the NEF, concepts may be ‘bound’ together and ‘combined’ within a scene using the mathematical operations of circular convolution (\*) and vector addition (+).

Consider an example making use of the subject-verb-object structure. Given a vocabulary of neurally representable concepts

{CAT, DOG, RED, TALL, SMELL, SUBJECT, VERB, OBJECT},

they may be used to describe the scene:

SMELL\*VERB + TALL\*CAT\*SUBJECT + RED\*DOG\*OBJECT

representing a tall cat smelling a red dog. Note that the order of binding and combining does not matter due to the associative and commutative properties of convolution and addition. Querying for the object of the scene can be done by convolution with the inverse (i.e. ‘unbinding’) of *OBJECT*:

$OBJECT^{-1} * (TALL*CAT*SUBJECT + SMELL*VERB + RED*DOG*OBJECT)$

and would return in the neural representation a vector close to *RED\*DOG*.

**Model and discussion**

The model consists of four subnetworks implemented in the neural simulator Nengo: (1) associative memory, (2) working memory, (3) basal ganglia, (4) thalamus. The English alphabet is encoded into associative memory and only

retrieved into working memory when given a cue to recite that is processed by the basal ganglia and routed through the thalamus. Figure 1 presents a diagram of the connections between different subnetworks.

The associative memory network encodes the sequential pattern of the alphabet. The SPA vocabulary is defined to be the 26 letters of the English alphabet and an additional concept for prompting recitation. Encoded in the heteroassociative memory is the relationship of the first 25 letters to their respective following letter, as well as the relationship of prompting recitation to the first letter *A*. The subnetwork is configured to be a winner-take-all network: given a semantic pointer, only one associated semantic pointer is recalled (i.e.  $PROMPT \rightarrow A, A \rightarrow B, B \rightarrow C, \dots, Y \rightarrow Z$ ).

The working memory network is a simple neural population with a recurrent connection that acts as a feedback loop such that the desired target semantic pointer is retained. Mathematically, this is equivalent to the high-dimensional integration over time of a vector (representing a semantic pointer) with a control that gates input signals.

The basal ganglia network is responsible for winner-take-all action selection and consists of component subnetworks: striatum, subthalamic nucleus, globus pallidus internal, globus pallidus external, substantia nigra pars reticulata. The selected action is based on a context-dependent utility signal for each action. Two actions are defined for the model:

**Recite** If working memory is neurally representing any letter from the alphabet then working memory begins to recall from associative memory the next letter.

**End** If working memory is neurally representing the letter *Z* then working memory begins to forget everything and resets.

The thalamus network routes the output of the basal ganglia to working memory. The firing within the thalamus represents the desired action selected and propagates the signal forward.

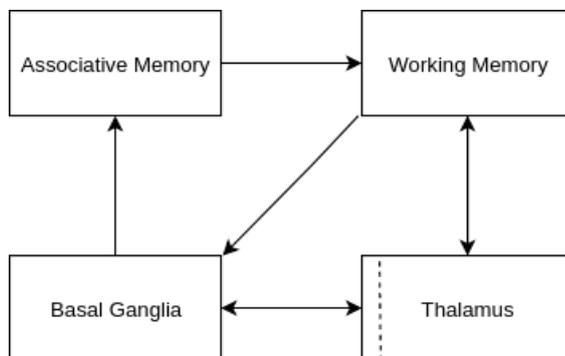


Figure 1. Model connectivity between subnetworks. Note that the signal towards the dashed line is attenuated under simulated lesions.

Lesions are simulated by reducing the neurons within the thalamus from 3800 to 3010. The number of neurons in both associative memory, working memory, and basal ganglia is fixed at 1300, 3200 and 4200, respectively. Figure 2 presents semantic pointer activation (representing letters) within working memory over simulation time for both unlesioned and lesioned recitations. The repetition of light blue ('F'), blue ('G'), green ('H') and orange ('I') peeks seen in the right graph show that the model incorrectly recites the alphabet in a structured way: the order of letters in the alphabet are preserved before continuing correctly.

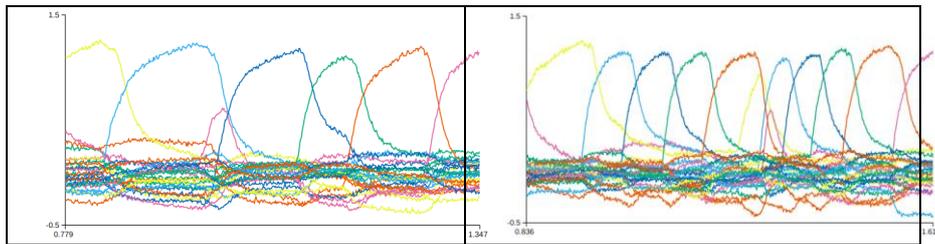


Figure 2. Recitation with (right) and without (left) simulated lesion. Without lesion: 'EFGHIJ'. With lesion: 'EFGHIFGHIJ'.

## Conclusion and future directions

This exploratory experiment suggests that spiking neural networks may be useful in modelling neurodegeneration by reducing the size of neural populations within subnetworks to simulate lesions. Additionally, the use of semantic pointers to represent higher-order concepts within a rigorous mathematical framework allows for the modelling of cognitive tasks. Further directions include implementing subnetworks that better represent regions of the brain in order to produce more biologically plausible simulations and to model language within the Semantic Pointer Architecture.

## References

- Eliasmith, C., Anderson, C.H. 2004. *Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems*. MIT Press.
- Eliasmith, C. 2013. *How to build a brain: A neural architecture for biological cognition*. Oxford University Press.
- Stewart, T.C., Tripp, B., Eliasmith, C. Python scripting in the Nengo simulator. *Frontiers in neuroinformatics* 3, 7.

# Emotional prosody perception in Italian as a second language

Emanuela Paone<sup>1</sup>, Manuela Frontera<sup>2</sup>

<sup>1</sup>Department of Humanities, University of Calabria, Italy

<sup>2</sup>Department of Cultures, Education and Society University of Calabria, Italy

DOI:

## Abstract

The present research aims at investigating the recognition and perception of emotional prosody in L2. 10 Russian learners of Italian and 33 native speakers of Italian (acting as a control group) were involved in a perception experiment. The test consisted of 8 stimuli, 4 in Italian and 4 in Russian, conveying different emotions. Participants were asked to listen to the stimulus, to recognize the emotion and to judge the intensity of the emotion on a scale from 0 to 3. Results showed that both groups were able to identify stimuli in both languages at above chance level, nevertheless they decoded the stimuli expressed in their L1 with higher accuracy. Furthermore, the two groups rated the intensity of Italian stimuli in a similar way, but differed significantly in judging the intensity of Russian stimuli.

Key words: Second Language Acquisition, emotional prosody perception, Italian L2

## Introduction

Some studies provide evidence that L2 learners from "distant" cultures experience increasing difficulties in "identifying emotion in the target language and in judging the intensity of that emotion than do fellow learners from "closer" cultures with similar levels of proficiency" (Dewaele 2005: 376; De Marco, Soriano, Paone 2017). Other studies suggest that L2 skills may sometimes interfere with emotion recognition from speech prosody (Bhatara, Laukka, Boll-Avetisyan, Granjon, Elfenbein, Bänziger 2016).

In light of the above, the present research aims at investigating the perception of emotional prosody by learners of Italian as a second language coming from a "distant" culture. We will address the following research questions: first, are learners of Italian with an upper-intermediate level of proficiency (B2) able to recognize emotional utterances in the L2 and to rate the emotional intensity as well as native speakers do? Second, is there an in-group advantage effect (Elfenbein, Ambady 2002), i.e. is participants' decoding ability more accurate when emotions are expressed in their L1? To give an answer to these questions native speakers of Italian and L2 Italian learners coming from Russia were involved in a perception experiment, to test their ability to decode emotions expressed in both languages (Italian and Russian).

## Method

### Participants

10 learners of Italian (6 F and 4 M, mean age= 25, age-range = 22-30) coming from Russia with B2 level of proficiency (assessed by a test), living in Italy since 18 months and attending Italian courses at the University of Calabria, and 33 native speakers of Italian (23 F and 10 M, mean age= 33, age-range 16-35), coming from South Italy, with no competence in Russian, completed the task.

### Stimuli

Participants were presented with utterances spoken in Italian and Russian produced with 3 different emotions (*cold anger, sadness, joy*) and without any emotional colouring (*neutral speech*). The standard sentences were "Lo hanno portato" ("they brought it") in Italian, and "Это Привезли" in Russian (collected and validated in previous research, see Paone 2017). The test consisted of 8 stimuli (4 in each language), conveying the three emotions and *neutral speech*.

### Procedure

Stimuli were presented using an online survey tool. Participants were asked to listen to the stimulus and to recognize the emotion expressed choosing from nine options (*anger, disgust, joy, fear, surprise, sadness, neutral, other*) and to judge the intensity of that emotion on a scale from 0 to 3. Learners were given instructions both in Italian and in Russian. The response options included different basic emotions and the option "other" in order to increase the reliability of the test, so the chances of random recognition were low, given the large number of possible answers. The emotional utterances were played randomly; no time limit was imposed.

Data from the perception experiment was analysed, taking into account correct identifications of stimuli in Italian and in Russian for each group. Binomial tests were performed for each stimulus type to determine whether they were correctly identified more often than chance. Repeated measures ANOVA and post-hoc tests were carried out in order to verify whether the two groups of participants differed in recognition accuracy and in the stimulus intensity evaluation.

## Results

In a multiple-choice test with 7 options, the probability of chance guessing is around 14,3%. From the results of the binomial tests, it appears that the global percentages of correct identifications for each emotion scored by both groups are above chance level ( $p < 0,05$ ). This means that the three emotions

and the neutral speech were correctly discriminated irrespective of the language in which the utterance was spoken. This result is in line with previous studies that claimed the ability of human beings to decode emotional speech even in a language other than their own.

Figure 1 shows the percentages of correct responses scored by both groups with regard to the stimuli in Italian and in Russian. As regards stimuli in Italian (histogram on the left), Russian learners were able to identify *sadness*, *anger* and *neutral speech* with the same accuracy as native speakers. Only the decoding of *joy* raised more difficulties, indeed the learners scored a lower percentage than native speakers' (46,6% vs. 70,7%), moreover the confusion matrix revealed that the major confusions occurred with *neutral speech* and with *surprise* in 16,6% of cases. The ANOVA results confirmed that there were no significant differences between the groups, but the stimulus types (*sadness*, *joy*, etc.) had an effect on the recognition scores. Post-hoc tests confirmed that NSs of Italian were better at discriminating the utterance of *joy* than Italian learners ( $p < 0,001$ ).

As regards stimuli in Russian (Figure 1, histogram on the right), Italians' accuracy level was lower than learners'. Indeed, significant differences were found in the case of *joy*, *sadness* and *neutral speech* ( $p < 0,001$ ), which were decoded with higher accuracy by Russian learners. The confusion matrix showed that Italians confused *sadness* with *neutral speech* in 16,1% of cases, and with *surprise* in 12,1%. *Joy* was confused with *surprise* (28,2%) and *fear* (21,2%).

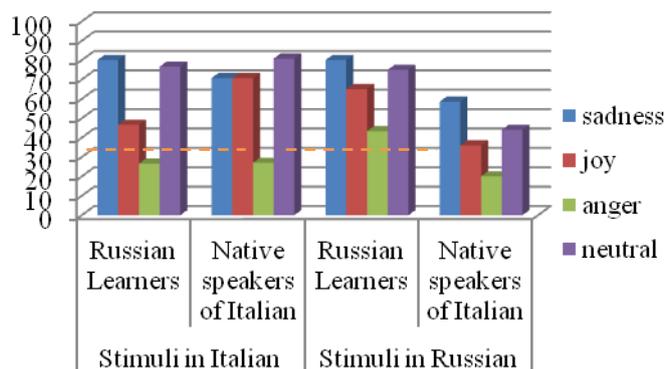


Figure 1. Percentages of correct responses scored by both groups with regard to the stimuli spoken in Italian (on the left) and in Russian (on the right). The dotted line represents the level of chance guessing (14,3%).

To establish the differences between the two groups in terms of emotional intensity, we took into account correct responses, excluding therefore erroneous recognition scores. A repeated measures ANOVA was carried out in order to determine the effects of the stimulus language (Italian- Russian) and stimulus type (*cold anger, joy, sadness, neutral speech*) on the intensity scores given by both groups. Results revealed that both variables had an effect. Post-hoc tests results showed that the two groups rated the Italian stimuli in a similar way ( $p > 0,05$ ), but differed significantly in judging the intensity of Russian stimuli ( $p < 0,05$ ). Indeed, native speakers of Italian perceived them as less intense than Russian learners did, especially in the case of *joy*.

## Conclusions

Russian learners and native speakers of Italian were able to identify emotions in Italian with similar levels of accuracy (except for *joy*), but Russian learners performed better in their L1, suggesting an in-group advantage effect. These results also suggest that learners' L2 knowledge might contribute to the ability to infer emotions from L2 speech prosody and to judge the intensity as well as native speakers do. Indeed, Italian natives scored lower percentages of correct identifications in Russian, a language they did not know at all, and judged the intensity differently from the learners. Further investigations are needed to clarify this point. These results might contribute to shed light on emotional prosody recognition and perception in a second language.

## References

- Bhatara, A., Laukka, P., Boll-Avetisyan, N., Granjon, L., Anger Elfenbein, H., & Bänziger, T. 2016. Second Language Ability and Emotional Prosody Perception. *PloS one*, 11, e0156855. doi:10.1371/journal.pone.0156855
- De Marco, A., Soriano, P., Paone, E. 2017. L'acquisizione delle emozioni nell'italiano non-nativo. In Gudmundson A., Alvarez López, L., Bardel, C. (eds.) 2017, *Languages: Multilingualism and Language Acquisition*, 57-86. Berna, Peter Lang.
- Dewaele, J.-M. 2005. Investigating the psychological and emotional dimensions in instructed language learning: obstacles and possibilities. *The Modern Language Journal* 89, 367–380.
- Elfenbein, H. A., Ambady, N. 2002. Is there an in-group advantage in emotion recognition? *Psychological Bulletin* 128, 243–249.
- Paone, E. 2017. La comunicazione delle emozioni in una lingua non nativa: un'indagine acustico-uditiva, Phd Thesis, Department of Political Science, University of Calabria.

# Experimental customisation of the Versioning Machine

Elsa Pereira<sup>1</sup>, Roman Bleier<sup>2</sup>

<sup>1</sup>CLUL, University of Lisbon, Portugal

<sup>2</sup>ZIM, University of Graz, Austria

DOI:

## Abstract

This paper deals with challenges in adapting the XML-TEI publishing framework Versioning Machine to compositional drafts of 20th-century literary works and describes the main customisations that have been implemented to suit a genetic edition of poetry by Pedro Homem de Mello. The case study emphasises that even minimal customisations require technical work that may go beyond an editor's skill.

Keywords: representation, drafts, TEI, Versioning Machine, customisation

## Representation of compositional drafts

Early 20th-century literary works are often documented in a variety of authorial witnesses, including draft manuscripts with several layers of additions, deletions, and rearrangements. A scholarly edition with a genetic orientation – aimed at achieving an insightful overview of the compositional and revisional development of those works – needs to represent the overwriting layers in each draft, while signalling the successive variants that occur across different witnesses.

In digital scholarly editing, the process involves two main stages: data modelling and presentation. The first stage is achieved using descriptive markup to produce machine-readable transcriptions, and while graph and hypergraph data models such as TAG (Dekker & Birnbaum 2017) are emerging as an alternative to the XML paradigm, the standard maintained by the Text Encoding Initiative [TEI] prevails as one of the most robust schemas available. It allows for a representation of both the materiality and the textual dimension of manuscripts and makes it possible to combine intra- and inter-documentary variation, using chapters 11 and 12 of the TEI P5 guidelines. The second stage converts the encoded texts into a visual layer to be accessed by the reader and is achieved by using XML transformation and query languages (XSLT, XQuery, XPath), as well as HTML, CSS, and JavaScript, to build an interactive graphical user interface. Such a wide range of technologies requires extensive technical support, which is not always available to individual or discretely funded projects.

The TEI community is aware of this problem and has developed tools and publishing solutions for XML-TEI, shared as open-source. Among the light-

weight solutions available, mention should be made of the TEI Boilerplate (Walsh et al.), the JavaScript library CETEIcean (Cayless & Viglianti 2018), EVT (Di Pietro & Rosselli Del Turco 2018), as well as the Versioning Machine [VM] (Schreibman et al. 2003; 2016) – one of the earliest and most widely known. Originally developed to trace the composition history of poems written by Thomas MacGreevy, the tool was used in several international projects and has been adapted for special publication needs over the years.<sup>1</sup>

### **Customisation of the Versioning Machine**

VM is specifically conceived to display multiple versions of text encoded with the TEI critical apparatus tag-set and additionally allows a limited number of elements to represent intra-documentary variation. By editing the tool's XSLT and CSS stylesheets,<sup>2</sup> we have achieved a wider range of presentational features, to suit an edition of print material and composition drafts by the Portuguese poet Pedro Homem de Mello (1904-1984). The alterations concern the display of three main categories: metadata, draft layers of revision, and inter-documentary variation.

The default VM setup provides a basic presentation of metadata, contained by the standard elements of the `tei:teiHeader`, to be displayed in the so-called “Bibliographic Panel”. However, an edition for the genetic orientation, focused “on the documents as sources of evidence of textual development and change through time” (Van Hulle & Shillingsburg 2015: 36), requires more detailed descriptions. For this reason, we included a “Compositional Synopsis”, containing the information from the `tei:creation` element (encoded within the *profile description*) and modified the “Original Source” section, by adding and structuring the content of `tei:msDesc` and `tei:biblStruct` (within the *source description*), as illustrated by example 1.<sup>3</sup>

Several alterations were made to suit the multi-layered transcription of complex documentary sources. Alternative readings, for instance, are not covered by the default VM setup, but it was necessary to include them for the project, since draft manuscripts occasionally have two or more phrasings at specific points of the text, without showing the author's preference for any of the options listed. In the XML-TEI this may be encoded using the `tei:seg` and `tei:add` elements with an `@type="alternative"`, for which we had to create a transformation rule, displaying one reading above the other, as seen in example 2.

Also important was rendering encoded gaps and significant spaces, as well as extending the display of authorial additions, through the processing of the values "top", "bottom", "marginLeft", "marginRight", "overleaf", and "opposite" of `@place`, since VM only deals with "above" and "below" by default (see example 3). However, revisions that involve several lines or

stanzas present a challenge. If the `tei:add` and `tei:del` elements are not allowed, due to conflicting XML hierarchies,<sup>4</sup> the TEI guidelines suggest using *milestone* or *empty elements*: i.e. `tei:addSpan` and `tei:delSpan` with a `@spanTo` attribute pointing to the ID of a `tei:anchor` that indicates how far the addition or deletion goes. The default stylesheets of VM do not contain instructions to process these milestone-delimited ranges, so new transformation rules were required (see example 4).

Another specificity of draft manuscripts is the repetition of revised stanzas or lines to clarify the wording. The TEI guidelines suggest encoding such clarifications with the `tei:retrace` element, which was introduced in 2011, as part of an encoding model for genetic editing. VM was designed for the critical apparatus tag-set and, therefore, does not support this element by default. A template rule had to be implemented to process and display `tei:retrace` and its `@rend` attributes, as shown in example 5.

For the same reason, VM does not deal with displacements of text (indicated by arrows or other symbols in the author's drafts), and new transformation rules were necessary to process these occurrences. In the project's encoding model, `tei:div` containing displaced text is specified by an `@type` with the value "displacedFragment", a `tei:metamark` indicates the place of insertion, and both elements are linked using an ID. Example 6 illustrates the rendering achieved by the modified stylesheets.

Besides the presentation of internal layers of revision, customisations were finally made to display the intricacies of inter-documentary variation in works with multiple drafts. According to the TEI P5 guidelines, apparatus entries may nest, using parallel segmentation with optional location-based referencing. While the VM is compatible with this encoding method, the default setup obscures the display of subvariants, because of the indistinct yellow highlight applied to all `tei:rdg` elements. We changed this by applying a different background colour whenever `tei:app` elements nest with different `@loc` ID, as illustrated by example 7.

## Conclusions

By looking at one specific software application that was developed for the publication of TEI data – the Versioning Machine – this paper explored the customisations necessary to adapt the software to an edition's needs and encoding model. In this case, a series of alterations had to be made to combine intra- and inter-documentary variation in poems where multiple drafts coexist.

The TEI tries to suit a diverse community of humanities scholars, which requires great flexibility to support different textual models and editing strategies. This circumstance means that the standard is not particularly well suited for the development of generic publication solutions. While some of

the changes described in this paper could be achieved with minimal changes to the VM source code, it still requires an understanding of web technologies to implement them, and not every textual scholar with basic XML knowledge has these skills or the resources to pay for them. Hence, the 1S/1P/1DH (one scholar, one project, one digital humanist) paradigm described by Robinson (2013) inevitably persists. The question that remains is how this can be overcome in the future.

## Notes

1. A selection of projects is listed on the website: <http://v-machine.org/vm-in-use>. The Wandering Jew's Chronicle is a good example of special customisation.
2. See `vmachine.xml` and `vmachine.css` in the "src" folder of our VM instance: [https://www.dropbox.com/sh/etrhbpwahu83j1/AAAE7nWx9woNrdF5Hu\\_aU41na?dl=0](https://www.dropbox.com/sh/etrhbpwahu83j1/AAAE7nWx9woNrdF5Hu_aU41na?dl=0)
3. Examples mentioned in this paper are in the "samples" folder of our VM instance.
4. See chapter 20 "Non-hierarchical Structures" of the TEI guidelines.

## Acknowledgements

The authors wish to thank the VM team for developing the Versioning Machine and Gael Vaamonde for his help during customising experiments.

## References

- Cayless, H., Vigiante, R. 2018. CETEIcean: TEI in the Browser. In Proceedings of Balisage: The Markup Conference 2018. Balisage Series on Markup Technologies 21. <https://doi.org/10.4242/BalisageVol21.Cayless01>.
- Dekker, R.H., Birnbaum, D. J. 2017. It's more than just overlap: Text As Graph. In Proceedings of Balisage: The Markup Conference 2017. Balisage Series on Markup Technologies 19. <https://doi.org/10.4242/BalisageVol19.Dekker01>.
- Di Pietro, C., Rosselli Del Turco, R. 2018. Between Innovation and Conservation: The Narrow Path of User Interface Design for Digital Scholarly Editions. In Bleier, R. et al. (eds.) 2018, *Digital Scholarly Editions as Interfaces*, 133-163. Norderstedt: Books on Demand.
- Robinson, P. 2013. Why digital humanists should get out of textual scholarship. <http://scholarlydigitaleditions.blogspot.com/2013/07/why-digital-humanists-should-get-out-of.html>.
- Schreibman, S., Kumar, A., McDonald, J. 2003. The Versioning Machine. *Literary and Linguistic Computing* 18-1, 101-107. <https://doi.org/10.1093/lc/18.1.101>.
- Schreibman, S., et al. 2016. Versioning Machine 5.0. <http://v-machine.org>.
- Text Encoding Initiative. 2019. P5: Guidelines for Electronic Text Encoding and Interchange. <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>.
- Van Hulle, D., Shillingsburg, P. 2015. Orientations to text, revisited. *Studies in Bibliography* 59, 27-44.
- Walsh, J. et al. TEI Boilerplate. <http://dcl.ils.indiana.edu/teibp/index.html>.

# Camilian lexical substitutions

Carlota Pimenta

Centre of Linguistics, University of Lisbon, Portugal

DOI:

## Abstract

With the theoretical and methodological framework of genetic criticism, this research focuses on the lexical substitutions made by Camilo Castelo Branco, one of the most relevant Portuguese Romantic writers, in the autographs of *Novelas do Minho* and *Amor de Perdição*, two major works of his. Through their analysis, an attempt is made to clarify his lexical ascertainment, specifically regarding lexical complexification, in two different periods of Camilian literary production.

Key words: genetic criticism, Camilo Castelo Branco, lexical substitutions

## Introduction

In my PhD thesis (Pimenta 2017), I transcribed, collected and sorted in a database the corrections found in the manuscript of *Novelas do Minho*, classified according to an analytical typology that revised and expanded the model previously proposed by Castro (2007). The present study now proposes to explore one of the types of genetic phenomenon registered at that time, the lexical substitutions, trying to demonstrate the usefulness of their analysis to deepen the knowledge about the way Camilo handled language.

This study explores the hypothesis proposed by Feijó (2011), according to which the continuous polishing of language by Camilo, often translated into a wider linguistic choice, is one example of continuity in Camilian writing. For this study, a sample of four novellas edited in successive years was selected – *Gracejos que matam (GM)* and *Cego de Landim (CL)* (1875), *Maria Moisés (MM)* (1876) and *A Viúva do Enforcado (VE)* (1877) – and the substitutions of the novel *Amor de Perdição (AP)* (1861) were also analysed (based on Castro 2007), in order to enable the comparison of this aspect in works of different periods.

## Method

The method was based on examining the substitutions that include less common terms. This method involves a margin of error that arises from the fact that nowadays we do not know the frequency of use that each word had at the time Camilo wrote. In an attempt to ascertain the degree of strangeness of the selected words, an exhaustive check of these terms was made in António de Morais Silva's *Dictionary of the Portuguese Language*, which

constitutes “the most important reference in the history of Portuguese lexicography” (Verdelho 2003). Camilo Castelo Branco used it frequently and advised his readers to consult it.

## Results

This study concluded that of 150 substitutions selected in *Amor de Perdição* and of 500 substitutions selected in *Novelas do Minho*, 20 less common words from the novel and 60 less common words from the novellas are not to be found in the second edition (1813) of Morais Silva's dictionary. This, for example, is the case with the noun “*objurgatória*” (censure) (*AP*: 20) and the adjective “*esparavonada*” (*GM*: 26), derived from the noun “*esparavão*”, a veterinary term which designates a tumor below the curve of a horse's leg.

The less common words that are missing in the 2<sup>nd</sup> edition were added in later editions. In fact, Morais Silva's dictionary, which appeared in 1789, was successively reprinted throughout the nineteenth century, with additions involving nomenclature and quotations from modern writers. Some words did not yet exist in the 6<sup>th</sup> edition (1858), the latest edition available on the dates Camilo wrote the novel and the novellas. This is, for example, the case with the term “*amantar*” (cover with a blanket), which was only added in the 7<sup>th</sup> edition (1878). This word appears in the sentence “*deixe-me ir <desaparelhar> amantar a egua*” (“let me <unsaddle> put a blanket on the mare”)<sup>1</sup> (*AP*: 132), that is the example presented in the 10<sup>th</sup> edition (1949).

Indeed, the 10<sup>th</sup> edition often presents examples taken from Camilo's own works and some of these examples are precisely the final versions resulting from lexical substitutions found in *Amor de Perdição* and *Novelas do Minho*. This is also the case with the noun “*aferro*” (vehement desire), included in the 2<sup>nd</sup> edition, which appears in the 10<sup>th</sup> edition with the example “<o amor> o <desejo de viver> *aferro* á vida que lhe foge” (“<the love> the <desire to live> vehement desire for life that runs from him”) (*AM*: 274) and is also the case with the verb “*exorcismar*” where the 2<sup>nd</sup> edition explicitly states “*exorcizar* (exorcising) is what one should say”. In the 10<sup>th</sup> edition, this term is referenced with the sentence “– Vai-te, vai-te, jacobino; cruces, deabo, cruces! – <exclamou> *exorcismou* a tia” (“- Go, go, Jacobin; crosses, devil, crosses! – <exclaimed> *exorcised* the aunt”) (*MM*: 4). Finally, this also happens with the adjective “*seva*” (ignoble), added in the 5<sup>th</sup> edition (1844) and attested in the 10<sup>th</sup> edition with the example “Era uma <pungente>[↑*seva*] calúnia” (“It was a <painful>[↑ignoble] calumny”) (*GM*: 34).

Even more interesting is the case of words that were only added in the 10<sup>th</sup> edition using examples from Camilo resulting from his revision work. This is the case of the verb “*paxalizar*”, derived from “*paxá*” (pascha; idler,

in the figurative sense), in the phrase “consignadas aos Cressos da rua do Ouvidor, <e aos nossos irmãos>[↑que *pachalisavam* nas chacaras da Tejuca” (“consigned to the wealthy from Ouvidor’s street, <and to our brothers> that *idled* in the ranches of Tejuca) (*CL*: 6) and the adverb “*enxundiosamente*” derived from “*enxúndia*” (fat), which occurs in “<engordando a po>[↑arredondando-se] tão *enxundiosamente* que parecia todo ele a barriga do gigante Typhéu” (“<fattening>[↑rounding out himself] so *full of fat* that all of him looked like the belly of the giant Tython”) (*VE* 1: 11). Both cases result from innovations in the language.

Words that were classified in the 2<sup>nd</sup> edition as ancient or little used forms were also found in Camilian substitutions. An example of a little used word is the noun “*latíbulo*” (hiding place) (*CL*: 16). Examples of ancient forms are the adjective “*afreimado*” (irate), in the substitution “– Vamos, q é tarde, vamos! – disse mt<sup>o</sup> <açodada>[↑*afreimada*]” (“– Come on, it's late, let's go! – she said <rushed>[↑*irate*]”) (*VE* 1: 20) and the verb “*prantar*”, an ancient and popular form of “*plantar*” (put), which appears in the phrase “Vais-te <por>[↑*prantar*] a pé, rapariga?” (“Are you going to stand up, girl?”) (*MM*: 33). These two examples resulting from substitutions were presented in the 10<sup>th</sup> edition.

In the process of replacing a more common term with a less common one, one can sometimes see hesitations in writing that seem to provide evidence of doubt when establishing the less common term. Thus, for example, in the phrase “<bebeu>[↑<g<ar>/ol\guelejou> *gargalaçou*] da borracha uma vez de vinho” (“<drank>[↑<gargled> *gargalaçou*] from the wineskin a large amount of wine”) (*MM*: 3), the verb “*beber*” (drink) was first replaced by the form “*garguerejou*”, which was amended to “*golguerejou*”. These two forms resemble the verbs “*gargarejou*” (to conserve a liquid in the throat without swallowing it) and “*gorgolejou*” (when drinking, to produce the noise of gargling). This hesitancy in writing seems to suggest Camilo's doubts as to the form to be established, having finally opted for a similar but less common verb “*gargalaçou*” (to drink, by inserting the neck of the vessel into the mouth). This verb was only added in the 7<sup>th</sup> edition of Morais Silva's dictionary.

Although the process of replacing a more common term with a less common one is dominant in the *corpus* studied, the opposite also happens. Sometimes, a less common term is rejected and replaced by a more common one. Examples of rejected terms are the verb “*presar*” (*VE* 2: 28), which the 2<sup>nd</sup> edition considers as an ancient form of “*apresar*” (capture), and the verb “*jornandeando*” (travelling) (*AP*: 200), which is added in the 8<sup>th</sup> edition (1889) as a modern term. Reference is also made to the word “*ovarina*” (natural or inhabitant of Ovar) (*AP*: 24), added in the 9<sup>th</sup> edition (no date), and to the forms “*repungia-se*” (deeply tormented) (*AP*: 278) and

“*despedaçadora*” (thing that shatters) (VE 1: 24), that only appear in the 10<sup>th</sup> edition.

### **Conclusion and future perspectives**

Although the work of lexical ascertainment is more intense in the manuscript of *Novelas do Minho* than in the manuscript of *Amor de Perdição* (cf. Pimenta 2017), this study suggests that the search for lexical complexification is present in both works, which, having been produced in different periods of the life of the writer, seem to indicate that this may be a constant throughout the creative process of Camilo.

Future developments could be made based on a computer application designed to verify the frequency of each word or phrase replaced in *corpora* including Portuguese texts of the second half of the nineteenth century. A platform that would provide free access not only to the forms within the Camilian linguistic heritage which are already known (originating from the published versions of his works) but also the linguistic forms that were rejected, as well as the genetic relation between them, would make a significant contribution to the study of the Camilian lexicon and its practical applications in literary and cultural analysis and would also provide a meaningful contribution to studies of Language History and Lexicography.

### **Notes**

Genetic symbols: < > deletion; [↑ ] substitution above the line; / \ overlapping substitution.

### **References**

- Castro, I. 2007. Camilo Castelo Branco, *Amor de Perdição*. Critical and genetic edition. Lisbon, IN-CM.
- Feijó, E. 2011. *O Legado do Último Camilo Romancista e a (Auto)Cilada realista*. Vila Nova de Famalicão, Casa de Camilo.
- Pimenta, C. 2017. *O processo de escrita camiliano em Novelas do Minho. Análise genética*. PhD thesis in Textual Criticism. School of Arts and Humanities of the University of Lisbon.
- Verdelho, T. 2003. *O dicionário de Morais Silva e o início da lexicografia moderna*. In *História da língua e história da gramática – actas do encontro*, 473-490, Braga, Universidade do Minho.

# Spatial anaphora in Hungarian

György Rákosi, Enikő Tóth

Department of English Linguistics, University of Debrecen, Hungary

DOI:

## Abstract

Hungarian employs reflexives in contexts of spatial anaphora to encode a referential dependency with clause-mate antecedents, and pronouns are generally unacceptable in the same configuration. This study shows that pronouns are also an option nevertheless in certain cases for most speakers, the major licensing factor being the distance between the antecedent and the figure located. Our results show that the constraints that govern the distribution of different types of pronominals in spatial anaphora in English are also relevant in Hungarian, albeit to a weaker degree.

Key words: anaphora, Hungarian, pronoun, PP, reflexive.

## Introduction

There is considerable variation both within and across languages in terms of the linguistic coding of spatial anaphora: personal pronouns and reflexive anaphors are in competition to establish a referential dependency with the local subject. In English, the default choice is the personal pronoun (1a), and the reflexive is only licensed under special circumstances (1b). These include the availability of antecedents whose point of view is described in the embedding discourse (Reinhart & Reuland 1993), and the description of locations which are close to the body of the antecedent (see Rooryck & Vanden Wyngaerd 2011 for an overview). So (1b), unlike (1a) strongly implies that the snake was close to the speaker's body.

In contrast, it is the reflexive which is acceptable across the board in these contexts in Hungarian (2b), and the use of the (*pro*-dropped) personal pronoun is a very marked option for most speakers, and some of the available linguistic literature simply considers the construction represented by (2a) ungrammatical (see, a.o., É. Kiss 1987: 185-186).

Rákosi (2012), however, reports on a questionnaire study where pronominal anaphora fared relatively well for most speakers in these contexts in first person (average ratings varying between 1 and 0 for individual examples on a 5 point-Likert scale from 2 to -2).

Our aim in this paper is to dig deeper into the nature of this variation. We restrict our attention to first person antecedents, as variation with some magnitude is expected only in non-third persons on the basis of the results of Rákosi (2012). We studied two types of PP constructions in a corpus study and a questionnaire study: PPs headed by postpositions that take non-case-

marked noun phrase complements, and PPs headed by postpositions which assign a specific case to their noun phrase complements. We found that the most important factor that facilitates the use of the pronoun in contexts of spatial anaphora is the nature of the space denoted: the pronoun is best (or even better than the reflexive in specific examples) when the space denoted is some distance away from the referent of the pronoun. Thus though the overall distribution of judgements is different, Hungarian does show sensitivity to a factor that is relevant in the case of English spatial anaphora.

- (1) a. I saw a snake near me.  
b. I saw a snake near myself.

(2) a.	%Láttam	egy kígyót	mellett-em.	Hungarian
	saw.1SG	a snake.ACC	near-1SG	
	‘I saw a snake near me.’			
b.	Láttam	egy kígyót	magam mellett.	
	saw.1SG	a snake.ACC	myself near	
	‘I saw a snake near myself.’			

### The corpus study and the questionnaire

To get a better grip on the design of our questionnaire, we conducted a corpus study first. The corpus study was performed on the Hungarian National Corpus (Oravecz et al. 2014), and it consisted of a manual selection of target examples including pronominal PPs with local antecedents in the clause. Such examples are not frequent, but they do occur in the corpus data. The ratio of locally coreferent pronominal PPs and reflexive PPs varies from 1:10 to 1:119 in our sample across the individual P types, reinforcing our claim that pronominal coding of spatial anaphora is an available option in Hungarian. We used the examples to construct test items for a pilot questionnaire. The aim of the questionnaire study was twofold. First, we wanted to check whether there is individual variation in the judgements concerning pronominal anaphora in Hungarian. It turned out that there is no significant variation, most subjects seemed to show comparable judgement patterns. Second, we wanted to have a baseline of ratings for our follow-up questionnaire.

The main questionnaire contained pairs of sentences, and the items in each pair were identical except for the choice of the PP-complement (pronoun vs reflexive, as in (2) above). Participants evaluated each sentence separately on a 5-point Likert scale from 1 to 5. Our underlying intention was to force the subjects to be aware of the competition between pronouns and reflexives. Since pronouns are a marked choice in this domain in Hungarian, if they still receive higher ratings in this specific comparison

task, then we have strong evidence that the pronominal coding is a grammatical option. Or it is so under special circumstances: our hypothesis was that pronouns will be acceptable if the location denoted by the PP is not close to the referent of the antecedent (being far is taken to be anywhere in the zone which is beyond arm's reach, see Kemmerer 1999 on this). Overall, in a mixed design we had two within-subjects factors with two levels each: pronoun type: pronoun vs. reflexive and location: near vs. far. We also had a between-subjects variable, P-type: one group of participants rated examples with postpositions that do not assign a case on their complements (95 subjects), and the other group saw test items only with postpositions that do (105 subjects).

### The results and discussion

As expected, there was a main effect of pronoun type; in general, reflexives received higher ratings ( $F(198, 1) = 164.56, p < 0.001, \eta^2 = 0.45$ ). A more interesting finding revealed a significant interaction between pronoun-type and location:  $F(1, 198) = 531.006, p < 0.001, \eta^2 = 0.73$ , which is represented in Figure 1. This indicates that pronouns were slightly preferred when the location is beyond arm's reach of the referent, while reflexives received higher ratings when the location denoted by the PP is within the peripersonal space around the referent of the antecedent. A test item from the questionnaire in (3) below illustrates the pronoun vs. reflexive contrast in the far condition, where *szemben* 'opposite' is a case assigning postposition.

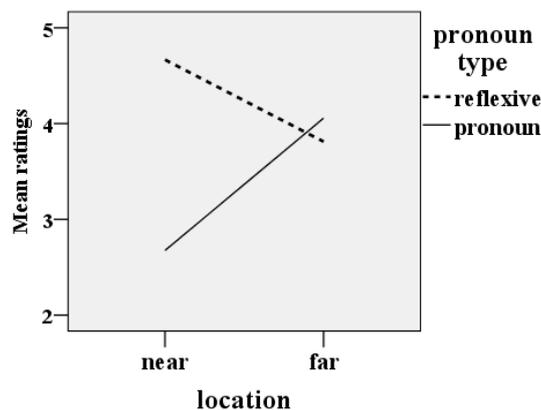


Figure 1. The interaction between pronoun type and location.

- (3) a. Köszöntöm itt            vel-em            szemben (...)    közgazdászt.  
       greet.1SG    here            with-1SG        near-1SG        economist.ACC  
       ‘I welcome the economist (...) here opposite me.’
- b. Köszöntöm itt            magam-mal       szemben (...)    közgazdászt.  
       greet.1SG    here            myself-with     opposite        economist.ACC  
       ‘I welcome the economist (...) here opposite me.’

(3a) with the pronoun received a mean rating of 4.66, while (3b), where the complement of the P is a reflexive got an average rating of 2.50. Hence, our hypothesis that pronouns might become a better option under special circumstances is supported by the results of the questionnaire. Analysing the significant three way interaction ( $F(1, 198) = 37.975$   $p < 0.001$ ,  $\eta^2 = 0.16$ ) shows that pronouns are more acceptable in examples with case-assigning Ps in the far condition than in the near condition. With Ps that do not assign case, pronouns in the far condition do not outperform reflexives, but they do represent a viable alternative since the average ratings are almost the same for pronouns and reflexives in this condition (4.02 and 4.27, respectively).

In sum, pronouns are an option in coding local spatial anaphora in Hungarian, but this option is restricted to non-third persons and to examples where the space denoted is further away from the body of the antecedent. This, to some extent, repeats the English pattern in a more restrictive setting.

## Acknowledgements

The Project no. 111918 (*New approaches in the description of the grammar of Hungarian pronominals*) has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the K funding scheme.

## References

- É. Kiss, K. 1987. *Configurationality in Hungarian*. Dordrecht, Reidel.
- Kemmerer, D. 1999. “Near” and “far” in language perception. *Cognition* 73, 35-63.
- Oravecz, Cs., Váradi, T., Sass, B. 2014. The Hungarian Gigaword Corpus. In: *Proceedings of LREC 2014*.
- Rákosi, Gy. 2012. A magyar birtokoszerkezetű viszonyjelölőkről, avagy kígyók szájfénnyel és esernyővel. [On possessive adpositional constructions in Hungarian, or snakes with lip glosses and umbrellas.] In É. Kiss, K., Hegedűs, A. (eds.) 2012, *Nyelvelmélet és dialektológia*, 154-173. Piliscsaba, PPKE BTK.
- Reinhart, T., Reuland, E. 1993. Reflexivity. *Linguistic Inquiry* 24 (4), 657–720.
- Rooryck, J., Vanden Wyngaerd, G. 2011. *Dissolving binding theory*. Oxford University Press.

# The acoustic characteristics of the Baraba-Tatar and Shor consonant ‘r’

Tatiana Ryzhikova<sup>1,2</sup>, Nikolay Urtegeshev<sup>1,2</sup>

<sup>1</sup>The Laboratory of Linguistic Anthropology, Tomsk State University, Russia

<sup>2</sup>Nadelyayev Laboratory of Experimental-Phonetic Researches, Institute of Philology, SB RAS, Russia

DOI:

## Abstract

The paper considers the acoustic characteristics of the Baraba-Tatar and Shor sound ‘r’. The experimental results show that it possesses a certain amount of noise and sometimes it is completely voiceless that makes it impossible to classify ‘r’ as a sonorant, but rather as a less-noise consonant. This term was proposed by V. M. Nadelyayev in 1960-s and fully describes the acoustic nature of the minority languages of Siberia and the Far East.

Key words: experimental phonetics, acoustic analysis of speech, Baraba-Tatar language, Shor language, less-noise consonants.

## Introduction

The problem of endangered languages preservation, investigation and analysis is one of the most acute nowadays. Because of the globalization processes minority languages are threatened by becoming extinct. Baraba-Tatar and Shor are two languages of the indigenous population of Southern Siberia. In the UNESCO Atlas of Endangered Languages they are both listed as severely endangered (Atlas 2010) and the linguistic situation becomes worse every year.

In 1960-1980-s V. M. Nadelyayev – the founder of the Laboratory of Experimental-Phonetic Researches of the Institute of Philology of the Siberian Branch of RAS (LEPR IP SB RAS) and the outstanding investigator and the leading specialist in the phonetics of the indigenous peoples' languages of Siberia and the Far East – noticed that the generally accepted rules of classifying consonants into noise and sonorant are inappropriate and unsuitable for the languages under consideration. Thus, he suggested that the consonants usually treated as sonorant should rather be called less-noise in the Siberian and Far-Eastern languages, because they do not comprise as much noise as noise consonants, but they are also not fully sonorant. In some positions, they can be manifested as fully or partially noise and moreover voiceless (Nadelyayev 1960).

The aim of this paper is to analyze the Barabian and Shor sound 'r' and show that it should be classified not as a sonorant but rather as a less-noise consonant.

### **Experiment**

The phonetic material has been collected in the places where the bearers of the two languages live: in the Baraba steppe and Mountain Shoria.

The Barabian recordings were made in 2000-2005 and 2012-2015. Presented here are the latest data. The number of speakers is 9, with their age differing from 22-25 up to 65-70 years. All of them speak Baraba-Tatar from birth, the second language is Russian, and some of them speak literary Tatar. The experimental program included words comprising sound 'r' in all possible positions in Barabian with the Russian translation. The subjects were asked to pronounce the target words three times in their native language. The recordings were made with a digital voice recorder HandyRecorderH4n with external noise filter. Processing of the linguistic material was done in SpeechAnalyzer. The voiced and voiceless components of the sound 'r' were distinguished and their relative length calculated.

The Shor material (the Verkh-Mrass (Upper-Mrass) subdialect) was also collected in the field in 2008. The second expedition to Shoria in 2016 showed that it was impossible to collect new material because of the age limits of the subjects or because of the ethnic changes in the villages population. Thus only material of 2008 is used in this paper. The speakers are the bearers of the Shor language (the Verkh-Mrass subdialect). Their number is 9. The age group is above 72 years old. All of them speak Shor as a native language, with Russian being the second language. The program was similar to the Barabian one and the process of recording material was the same. A digital voice recorder HandyRecorderH4n with external noise filter was used. Post-processing was done in the laboratory conditions in the SpeechAnalyzer Program. Target sound is 'r'.

### **Results and discussion**

In different languages sound 'r' has different articulatory-acoustic characteristics. The common trait for all of them is the presence of a short pause surrounded by a vocal context (sometimes there are two, three or more such sequences) and it is this combination that defines a vibrant consonant (Knyazev, Pozharitskaya 2011). In Baraba-Tatar 'r' is a forelingual vibrant (Ryzhikova 2004). In Shor it is a forelingual pharyngealized vibrant (Urtegeshev 2004). Russian 'r' is a forelingual voiced vibrant (Knyazev, Pozharitskaya 2011). Figure 1 shows the examples of the Barabian, Shor and Russian sound 'r' in the position between the vowels and in the word end.

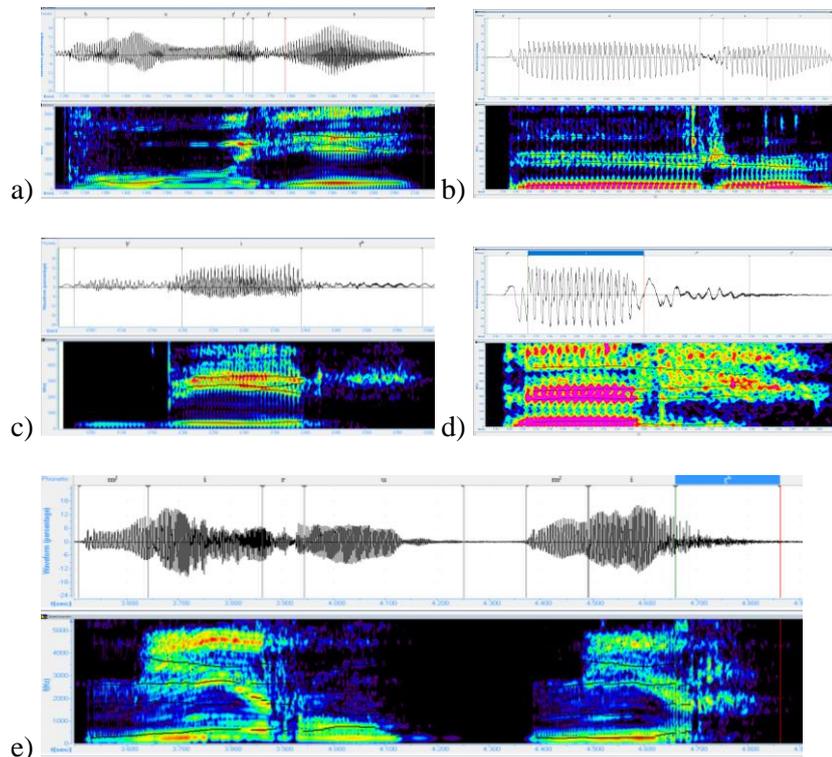


Figure 1. a) Barabian büre ‘a wolf’, b) Shor pörü ‘a wolf’, c) Barabian bir ‘give!’, d) Shor per ‘give!’, e) Russian miru mir! ‘peace in the world!’

The results of the acoustic analysis of the Baraba-Tatar and Shor languages turned out to be in agreement. The 'r'-type consonant in the Verkh-Mrass subdialect of Shor can be interpreted as less-noise, because it comprises noise component in all the positions in a word and in some cases it consists only of the noise part. Nevertheless, almost all realizations of the Shor phoneme /r/ can be interpreted as voiced ones, except the position in the very end of a word. Here the presence of voice depends on the vowel harmony of a word: if it is with front vowels, the allophones of /r/ are voiceless (the average amount of voice is 19-57 % of the sound length); if the word contains back vowels, voice comprises about 50-81 %.

In Baraba-Tatar, realizations of the phoneme /r/ also contain the noisy part, but the presence of the voice in their production does not seem to play an important role: even in the strongest position between two vowels the allophones of /r/ may be both voiced and voiceless. In the final position in the word the length of the realizations increases in 1,5 times of the average length of the sound and the number of voiceless components varies from 92 up to 100%. In general, the number of voiceless components in all

realizations of the phoneme /r/ is 71,3%. To sum up, it can be defined as less-noise with a different degree of voiced components.

In the Russian example, one can see a considerable amount of noisy components in the very end of a word and the absence of voice, though the bearers of the Russian language do not perceive it (Knyazev, Pozharitskaya 2011).

### **Conclusions**

To summarize, the investigation has proven that the classification of consonants at least in two Siberian Turkic languages into noise and less-noise units seems to be reasonable. Moreover, the presence and quantity of voice in the production of the phoneme /r/ in the languages under consideration vary significantly: in some positions the presence of voice is negligible in both of them. This fact allows us to assume that voice is not a constitutive-differential characteristic of the phoneme /r/ and it should be classified not as sonorant but rather as less-noise with the additional characteristic of being voiced or voiceless.

### **Acknowledgements**

The work was financially supported by the RF Government Grant № 14.Y26.31.0014 «Linguistic and ethno-cultural diversity of Southern Siberia in synchronic and diachronic aspects: interaction of languages and cultures».

### **References**

- Knyazev, S. V., Pozharitskaya, S. K., 2011. *Sovremenniy russkiy literaturny yazyk: Fonetika, orfoepiya, grafika i orfografiya*. Moscow, Akademicheskiy proekt, Gaudeamus. (Contemporary Russian literary language: Phonetics, orthoepy, graphics and spelling).
- Nadelyayev, V. M., 1960. *Proyekt universal'noy unifikirovannoy foneticheskoy transkripsii (UUPhT)*. Moscow-Leningrad. (The Project of the universal unified phonetic transcription).
- Ryzhikova, T. R., Urtegeshev, N. S., 2004. Less-noise forelingual consonant phonemes in Baraba Tartars' and Shor languages. In *Proc. 12th International Conference on Turkish Linguistics*, Izmir, North Cyprus.
- UNESCO Atlas of the World's Languages in Danger (<http://www.unesco.org/languages-atlas/> last consulted 27.07.2019).
- Urtegeshev, N. S., 2006. *Maloshumnye soglasnye shorskogo yazyka: artikulyatsionno-akusticheskiye osobennosti (mrasskiy dialect) in Alchikova, O. et al., Tyurkskiye yazyki: problemy i issledovaniya, 77-82, Gorno-Altaysk, Russia. (Less-noise consonants of the Shor language: articulatory-acoustic peculiarities (the Mrass dialect)).*

# Processing the meaning of English articles by Arab learners

Mona Sabir, Alaa Melebari

English Language Institute, King Abdulaziz University, Saudi Arabia

DOI:

## Abstract

This study investigates the question of whether L2 learners' online knowledge of specificity in English article choice mirrors their offline knowledge. Previous research found Hejazi-Arabic learners of English sensitive to the specificity feature when tested offline. In the current study, 68 Hejazi-Arabic participants and 23 native English speakers took an article elicitation task and a self-paced reading task. The article elicitation task results show learners' overuse of 'the' in the indefinite specific context. The self-paced reading task results show higher reaction times for L2 learners for the *correct* article choice in the indefinite specific context only compared to native speakers. This suggests that Arab learners encounter difficulty due to the form/meaning mismatch of [+specific] feature in this context.

Key words: processing, English articles, definiteness and specificity

## Introduction and theoretical background

The general findings in article acquisition research (see for example García Mayo & Hawkins, 2009) is that substitution errors often persist in L2 production even with advanced language learners. Ionin, Ko & Wexler (henceforth IKW) (2004) attribute substitution errors in contexts where there is a mismatch between the definiteness feature and the specificity feature (i.e. [−definite, +specific] & [+definite, −specific]) by learners from languages without articles (−ART) to the specificity feature which led to the formulation of the Fluctuation Hypothesis (FH). That is, (−ART) L2 learners assign articles based on specificity rather than definiteness. Later, Ionin, Zubizarreta and Philippov (2009) revised this FH prediction to only include overuse of *the* in the [−definite, +specific] context.

Previous research on Arabic learners of English (e.g. Almahboob, 2009 and Sabir, 2018) found that Arabic learners often overuse 'the' in the indefinite specific context when tested offline despite that Arabic is a (+ART) language. These studies, however, have only investigated Arab learners' offline knowledge of English articles. To the best of our knowledge, Arab learners' online processing of English articles have not been explored yet. In fact, studies that tested L2 learners' online processing of English articles, in general, are very limited. The closest study to our study is a self-paced reading that was conducted by Kim and Lakshmanan

(2008) who investigated the ability of intermediate and advanced Korean L2 learners of English to distinguish English articles based on specificity rather than definiteness. Their results showed that the control group and the advanced L2 learners fluctuate between the definiteness setting and the specificity setting while the intermediate L2 learners initially process articles based on specificity and then definiteness.

### **The current study**

In the current study, we investigate the question of whether Saudi-Hejazi<sup>1</sup> EFL learners' online knowledge of specificity in the English article system mirrors their offline knowledge. 68 intermediate-level Saudi-Hejazi Arabic speaking participants in addition to 23 native English speakers took an article elicitation task (AET/offline/explicit task) and a self-paced reading task (SPR/online/implicit task). The overall prediction for the current study is that if participants show sensitivity to specificity in the target context (the indefinite specific) compared to the other three contexts in terms of fluctuation (monitored by the AET) as well as longer latencies for the mismatch cases in this context (monitored by the SPR), then we can conclude that they have acquired the article system implicitly, but not explicitly. Contrastively, if they show sensitivity only in the AET, but not in the SPR, then the assumption is that the system has not been acquired explicitly nor implicitly.

### **Results**

For the AET, ANOVAs were conducted on the use of *the* vs. *a/an* for EFL learners and for the native control group. The results in table (1) show that native speakers performed as expected with no fluctuation, while definiteness is highly significant in EFL learners' article use. Additionally, specificity has a highly significant effect on learners' use of both *the* and *a/an*. This means that EFL learners are fluctuating between *the* and *a/an*.

For the SPR task, the ANOVA and the post-hoc tests reveal a difference in response times between the match and mismatch conditions for both native and non-native speakers. In general, the mismatch conditions yielded longer reaction times. This holds across all contexts except for the indefinite specific one where the effect is the opposite for non-native speakers, i.e. non-natives speakers took longer in the match condition compared to the mismatch condition. (Figure 1).

Table 1. Results of repeated measures ANOVAs for the natives and non-natives (\*\*p&lt;.001, \*p&lt;.01, \*p&lt;.05).

Factor	Use of the (natives)	Use of a (natives)	Use of the (non-natives)	Use of a (non-natives)
Definiteness	F (1, 22) = 5808***	F (1, 22) = 3480***	F (1, 407) = 25959***	F (1, 407) = 67353***
Specificity	F (1, 22) = .581	F (1, 22) = .042	F (1,407) = 50.826***	F (1, 407) = 310.136***
Definiteness x specificity	F (1, 22) = 9.85**	F (1, 22) = 2.77	F (1, 407) = 3.985*	F (1, 407) = 219.930***

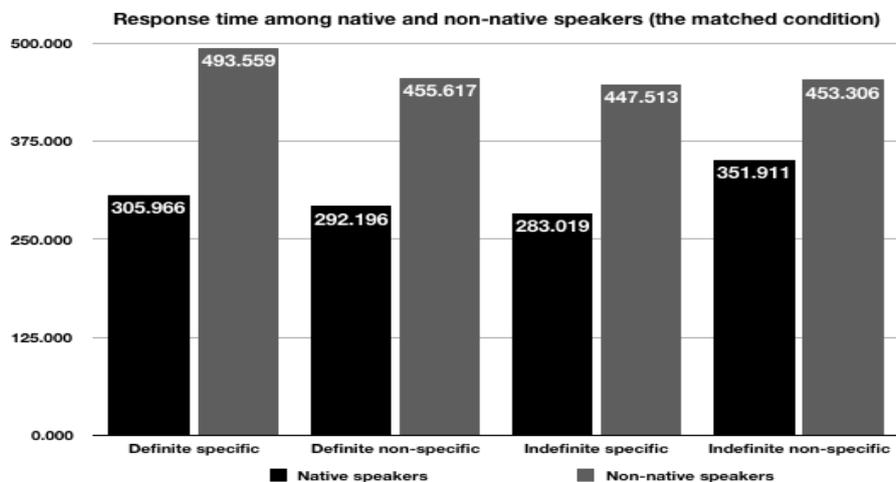


Figure 1. Response time for native and non-native speakers (match condition).

## Discussion and conclusion

The AET results show that Saudi Hejazi-Arabic learners of English fluctuate between the specificity and definiteness parameters, a finding that goes against García - Mayo, 2009 (among others) findings for L2 learners with a + ART L1. However, the SPR results show that despite that Arab learners generally take longer in the mismatch condition compared to the natives; we observe a significant switch in the indefinite specific context. The fact that their performance switches in the indefinite specific allows us to notice a specificity-over-definiteness processing preference. In other words, higher reaction times in the match condition of this context might be an instance of

confusion where they activate [+specific] reading and expect to see the article “the” but alternatively what they encounter is the article “a” which creates a mismatch between the activated [+specific] reading and the [–specific] article encountered. This suggests that they process articles, at least initially, based on specificity, which is consistent with Kim & Lakshmanan’s (2008) findings for Korean L2 learners whose L1 is a – ART language.

## Notes

Hejaz refers to the western region of Saudi Arabia (the western border on the Red Sea).

## References

- Almahboob, I. 2009. The L2 acquisition of English articles by L1 speakers of Saudi Arabic. (PhD), University of Essex.
- García-Mayo, M.d.P., Hawkins, R. (Eds.). 2009. *Second Language Acquisition of Articles* Amsterdam: John Benjamins Publishing Company.
- García-Mayo, M.d.P. 2009. Article choice in L2 English by Spanish speakers: Evidence for full transfer In M.d.P. Garcia-Mayo, R. Hawkins (Eds.), *Second Language Acquisition of Articles* (pp. 13-36). Amsterdam: John Benjamins Publishing Company.
- Ionin, T., Ko, H., Wexler, K. 2004. Article Semantics in L2 Acquisition: The role of specificity. *Language Acquisition*, 12(1), 3-69.
- Ionin, T., Zubizarreta, M., Philippov, V. 2009. Acquisition of article semantics by child and adult L2-English learners. *Bilingualism: Language and Cognition*, 18, 242-280.
- Kim, L.K, Lakshmanan, U. 2008. L2 article semantics and second language processing. In proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007), ed. Roumyana Slabakova et al., 103-117. Somerville, MA: Cascadia Proceedings Project. [www.lingref.com](http://www.lingref.com), document #1630
- Sabir, M. 2018. Specificity in article use: a theoretically informed classroom intervention. *Instructed Second Language Acquisition*, 2(2), 137-163.

# Neurodegenerative language impairments and transcranial stimulation

Clara Sanches<sup>1</sup>, Juliette Godard<sup>1</sup>, Marc Teichmann<sup>1,2</sup>, Antoni Valero-Cabré<sup>1</sup>

<sup>1</sup>Institut du Cerveau et de la Moelle Epinière, Sorbonne Université, France.

<sup>2</sup>Department of Neurology, Pitié Salpêtrière Hospital, France.

DOI:

## Abstract

The mental lexicon is the stock of words, in long-term memory, as an association of syntactic, semantic and phonological/orthographic structures. Being able to access word meaning rapidly and flexibly is essential for efficient communication and it involves highly specialized brain networks. Its breakdown is present in different neurodegenerative diseases, as Primary Progressive Aphasia (PPA), Frontotemporal dementia (FTD) and Progressive Supranuclear Palsy (PSP). Transcranial direct current stimulation can help improve language deficits in these diseases. We compared the language deficits of these patients with controls and look for the effects of tDCS in such deficits.

## Introduction

Language defines humans, and words are a distinctive language-specific part of human knowledge (Jackendoff 2002). Words are an association, in the long-term memory, of syntactic, semantic and phonological/orthographic structures, which is known as the mental lexicon (Jackendoff 2002). Research in the domain of language is often pursued at the word level because words are relatively well defined minimal units that carry many codes of analysis and processing distinctions (automatic vs. attentional). The access to the mental lexicon involves several mental representations that recruit temporal, parietal and frontal areas (Minicucci et al. 2013). The breakdown of critical nodes within this intricate system is present in different neurodegenerative diseases and at different levels.

Semantic and logopenic Primary Progressive Aphasia (svPPA, lvPPA) affect, respectively, the left anterior temporal lobe (ATL) and the left temporoparietal junction (TPJ) and impair conceptual knowledge (svPPA) and single-word retrieval (lvPPA) (Gorno-Tempini et al., 2011). The behavioral variant of Frontotemporal dementia (bvFTD) damages the prefrontal cortex (Rascovsky et al. 2011) and Progressive Supranuclear Palsy (PSP) affects subcortical structures but also the dorsolateral prefrontal cortex (DLPFC) (Paviour et al. 2006). Both impair language initiation and research mechanisms in the mental lexicon (Paviour et al. 2006).

Transcranial direct current stimulation (tDCS) is a non-invasive brain stimulation technique able to modulate cortical activity and promote

neuroplasticity and has been employed in the language domain with encouraging results (e.g. Cotelli et al. 2014). It operates by delivering a very weak constant current (1-2 mA) between at least two electrodes (an anode and a cathode) placed on separated scalp locations of the head. Single sessions of tDCS induce short lasting reversible modulations of cortical activity, whereas the accrual of periodical sessions enables longer lasting impacts by engaging neuroplasticity.

In this study we applied single sessions of tDCS to groups of patients with different neurodegenerative diseases to evaluate its ability to modulate language and its therapeutic potential. We also expected to contribute to a better characterization of the extent of the language deficits and the anatomical systems subtending them in such diseases.

## Materials and methods

Four groups of neurodegenerative patients: svPPA, lvPPA, bvFTD, PSP (n=12 for each) were recruited at the Pitié-Salpêtrière Hospital, Paris. Patients underwent 3 tDCS sessions (20 minutes, 1.59 mA, 0.06 mA/cm<sup>2</sup>) and immediately before and after each session they performed a series of language tasks. Each group was stimulated over a left cortical damaged region (anodal tDCS), over its right homologue (cathodal tDCS): ATL for svPPA, TPJ for lvPPA and DLPFC for bvFTD and PSP, and with sham tDCS. Four experimental tasks were used to characterize language impairments and to check for the effects of stimulation: *Letter fluency (LF)*, *Category decision (CD)*, *Semantic association (SA)* and *Picture naming (PN)*. Two versions for each task were designed and the order of application was counterbalanced between each group of patients, as well as the order of the three stimulation sessions. Fifteen healthy controls (HC) were also recruited to have normative values for each task. Healthy controls were not stimulated.

## Results

To characterize language impairments for each group, ANOVAs were performed for each task, with ‘group’ as the independent variable and ‘performance’ (before stimulation) as the dependent variable, with Tukey posthoc tests. Results are summarized in Figure 1.

Biophysically inspired computational models, using a template head, to predict tDCS current magnitude and distribution in the brain for each group can be seen in Figure 2.

To check for the effects of tDCS, ANOVAs contrasted pre- and post-stimulation performances by comparing left-anodal and right-cathodal with sham stimulation. Results can be found in Figure 3.

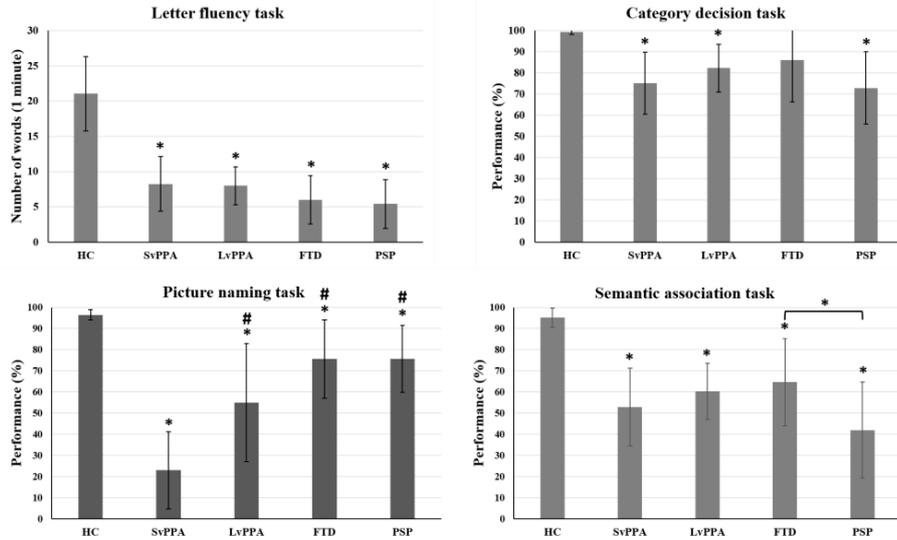


Figure 1. Group performance outcomes at baseline. \* significantly different from HC ( $p < 0.001$ ); # significantly different from svPPA ( $p < 0.05$ ); \* *with a claudator*: significantly different between them ( $p < 0.05$ ).

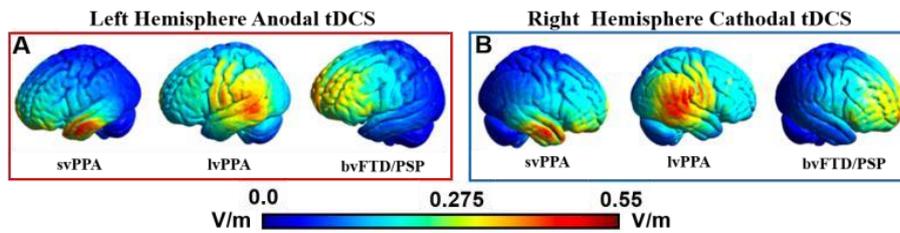


Figure 2. Computational models of current magnitude and distribution in the cortex. A) Anodal stimulation, left hemisphere view: current flows in, promotes neuron activity by reducing resting membrane potential (depolarization). B) Cathodal stimulation, right hemisphere view: current flows out, inhibits neuron activity by increasing resting membrane potential (hyperpolarization). Colour bar: signals the norm of the electric field at the cortical surface (V/m).

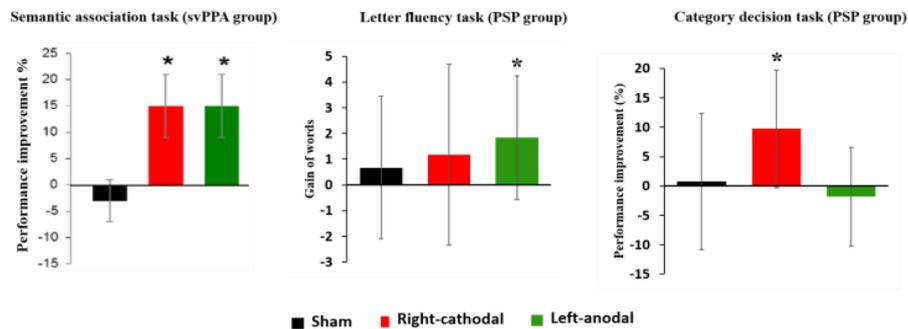


Figure 3. A) Improvements after both types of stimulation in the SA task for the svPPA group. B) Improvements after left-anodal stimulation in the LF task for the PSP group. C) Improvements after right-cathodal stimulation in the CD task for the PSP group. No other effects were found.

## Conclusion

All patients present language deficits in all different tasks, showing that language deficits extend beyond the expected ones for each group according to the localization of anatomical damage. TDCS induced currents in the brain that encompass the intended targeted region, however, with different behavioral effects for each group. We conclude that tDCS can prove beneficial in helping to contain deficits in neurodegenerative diseases such as those presented here.

## References

- Cotelli, M., Manenti, R., Petesi, M., Brambilla, M. et al. 2014. Treatment of Primary Progressive Aphasias by Transcranial Direct Current Stimulation Combined with Language Training. *Journal of Alzheimer's Disease*, 39, 799–808.
- Gorno-Tempini, M.L., Hillis, A.E., Weintraub, S., Kertesz et al. 2011. Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006-1014.
- Jackendoff R. 2002. *Foundations of Language*. (1st Ed.). New York, NY: OUP.
- Minicucci D., Guedish S., Blumstein S.E. 2013. An fMRI examination of the effects of acousticphonetic and lexical competition on access to the lexical-semantic network. *Neuropsychologia*, 51(10), 1980-1988.
- Paviour D.C., Price S.L., Jahanshahi M., Lees A.J., Fox N.C. 2006. Regional Brain Volumes Distinguish PSP, MSA-P, and PD: MRI-Based Clinico-radiological Correlations. *Mov Disord* 2006 ; 21(7) : 989-996 .
- Rascovsky K., Hodges J.R., Knopman D., Mendez M.F. et al. 2011. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 2011; 134(9); 2456-2477.

# Lexicalized pauses in Italian

Loredana Schettino, Violetta Cataldo

Department of Humanities Studies, University of Salerno, Italy

DOI:

## Abstract

This research explores lexicalized filled pauses in Italian tourist guides' speech, addressing the existing correlation between their formal, mainly phonetic, features and pragmatic functions. The study reveals that pauses'  $F_0$  contour correlates with their main functions, allowing to distinguish pauses working as controlled focusing devices, from more explicitly unintentional hesitation pauses.

Key words: Speech, Disfluencies, Lexicalized Filled Pauses, Discourse Markers

## Introduction

The spontaneous human speech comes with a number of phenomena affecting its 'fluency', namely pauses, repairs, repetitions, whose incidence may vary depending on the type of speech and its planning degree. It is by now recognized that some of such 'disfluencies' are charged with communicative values. In particular, pauses provide the listener with valuable meta-information about the ongoing speech (e.g. about discourse planning, its structure, speakers' modal disposition). Though, there is no evidence of speakers' control on their production (see Eklund 2004, Corley and Stewart 2008). Two forms of disfluencies are usually identified: repairs, something in the utterance is retraced and altered; hesitations, as temporary interruptions and delays in flowing speech (Lickley 2015). Hesitation pauses may be subdivided in unfilled (silent) pauses and filled pauses. Among filled pauses, those with a proper lexical form, here referred to as 'lexicalized' filled pauses (such as discourse markers), may be distinguished from the 'non-lexicalized' filled pauses (prolongations, vocalizations and the like).

This exploratory study aims to investigate disfluencies falling into the set of pauses. It follows a previous investigation on Silent Pauses (SP), Filled Pauses (FP, such as 'ehm' or 'eeh'), Vocalized Filled Pauses (VFP, segmental prolongations) (see Cataldo *et al.* in press) and extends the object of study to Discourse Marker Filled Pauses (DMFP, such as *diciamo* 'let's say', *quindi* 'so'). DMFPs are here identified as polyfunctional lexical elements, external to propositional content, which, given their semantic meaning, develop pragmatic functions due to the context they are uttered in (Bazzanella 2006).

To this end, the existing correlation between DMFPs' formal, mainly phonetic, features and pragmatic functions is addressed. Such studies might find application in automatic speech processing (see Origlia *et al.* 2018).

## Methodology

For this study, a dataset was considered of the main audio-visual corpus of visits led by expert guides in San Martino's Charterhouse (Naples). The dataset consists of circa 35' semi-spontaneous speech by three female tourist guides (G01, G02, G03). The selected data were annotated using ELAN, for a multilevel annotation of disfluencies, and PRAAT for the description and analysis of their phonetic realization.

DMFPs were analyzed according to the following parameters:

*Pragmatic functions* – 1) Interactional Function (IF), related to discourse planning within the interaction; 2) Meta-Textual Function (MTF), related to structuring and emphasizing information in speech; 3) Cognitive Function (CF), dealing with ongoing cognitive processes related to the propositional content and the illocutionary force (based on Bazzanella 2006)<sup>1</sup>.

*Phonetic features* – DMFP F<sub>0</sub> profile (rising, falling, flat, valley, peak); mean pitch in Hz (of the Tonal Unit, TU, and of the DMFP occurrence); creaky vocalizations; DMFPs distribution within the TU.

## Results

The incidence of the 639 DMFPs on the guides' speech is quite even across the speakers (one item uttered per about 9 words and about 17 per minute).

*Pragmatic functions.* Figure 1 shows DMFPs are mainly adopted as focusing devices, textual markers and/or boosters. In fact, out of 463 focusing devices, about the half is also boosters and a quarter carries out a text marker function as well. Despite this common tendency, more specific individual preferences reflect the idiosyncrasies found in the previous study.

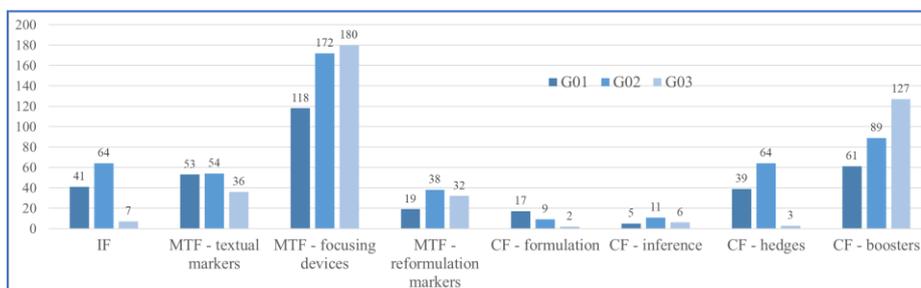


Figure 1. DMFP occurrences per speaker and per function.

G01 prefers an ‘on the fly’ strategy, employing a number of formulation pauses and hedges, fillers, and phatic devices; G02 tends to avoid non-lexicalized pauses altogether through ‘juxtaposition’ of utterances and uses a high number of lexicalized pauses instead; G03 adopts a more controlled, ‘rhetorical’ style, using DMFPs as focusing devices and boosters mainly.

*Phonetic features.* As shown in Table 2A, most often DMFPs are realized with a flat  $F_0$  profile or as a tonal trough within the TU (31%, 25%). Most noteworthy is the recurring  $F_0$  ‘peak’ realization (30%), which appears to correlate with focusing devices<sup>2</sup>. Considering focusing DMFPs (foc) as compared with the remaining DMFPs (non-foc.), a generalized linear mixed model testing the correlation between function (foc vs. non-foc) and  $F_0$  profile showed that the latter is a significant predictor of the function ( $\chi^2=20.74$ ,  $p<.001$ ). Pairwise comparison among fixed levels then showed that ‘peak’ and ‘valley’ are significant predictors of, respectively, foc. and non-foc. function.

Such results confirmed also the mean pitch values registered for both DMFPs and their TU (Table 2B). Only the 194 DMFP peak occurrences are realized with mean pitch values about 25 Hz higher than those of their TU. Standard deviation high values depend on the inter-speaker variability (the maximum excursion in G03 is double the one in G02) and on the fact that the ‘peak’ doesn’t always reach the highest point in the TU.

Table 2. DMFP occurrences and percentages per  $F_0$  profile and per speaker (A); mean pitch (Hz) of TUs and DMFPs,  $F_0$  difference between DMFP peaks and TU mean pitch (Hz) and its standard deviation values (B).

A	F <sub>0</sub> patterns					B	Mean pitch (Hz)			
	flat	valley	rising	falling	peak		tot DMFPs		peaks DMFP	
G01	84	33	2	7	47					
G02	96	79	26	25	47					
G03	19	47	4	21	100	G01	219,8	222,9	25,5	29,8
n. occ.	199	159	32	53	194	G02	214,6	212	24,5	18,8
% occ.	31%	25%	5%	8%	30%	G03	215,3	223,5	23,9	35,5

## Discussion and conclusions

As for the pragmatic level, DMFPs are found to diverge from FPs and VFps. Given their multifunctional nature, the former set is highly used to perform meta-textual functions (most often as focusing devices) and cognitive functions (as mitigating or boosting devices), unlike non-lexicalized pauses, mostly used as a tool of the speech planning process.

As for their phonetic features, DMFPs occur with a wider range of possible realizations and show more modulated intonation patterns, as they

have more segmental content, than the other voiced pauses. Furthermore, the ‘peak’ realizations, never found neither for FPs nor for VFPs, seem to distinguish structuring focusing DMFPs, working as deliberate rhetoric devices, from the other more explicitly unintentional lexicalized pauses (fillers and modal devices) showing a rather flat or valley pitch contour.

## Notes

1. IF, i.e. speaker-oriented devices (turn-taking, fillers, attention-getting, phatic, checking comprehension, requesting agreement/confirmation, yielding the turn) and addressee-oriented devices (interrupting, back-channels, attention confirmed, phatic, acknowledgement, agreement/confirmation); MTF, i.e. textual markers (introduction, transition, list, digression, ending, quotation), focusing devices and reformulation markers (paraphrase, correction, exemplification); CF, i.e. markers of formulation/inference and modulation devices (hedges and boosters).
2. Ex: *i principi ispiratori della comunità sono legati a quello che è un altro ordine esistente in Europa ed è l'ordine benedettino* - ‘The community’s guiding principles are tied to **what it is** another existing order in Europe that is the Benedictine Order’.

## Acknowledgements

Work funded by the Italian National Project PRIN ‘Cultural Heritage Resources Orienting Multimodal Experiences (CHROME)’ (#B52F15000450001).

## References

- Bazzanella, C. 2006. Discourse Markers in Italian: towards a ‘compositional’ meaning. In Fischered, K. (Ed.). *Approaches to discourse particles*, 449-464. Amsterdam, Elsevier.
- Cataldo, V., Schettino, L., Savy, R., Poggi, I., Origlia, A., Ansani, A., Sessa, I., Chiera, A. in press. Phonetic and functional features of pauses, and concurrent gestures, in tourist guides’ speech. *Studi AISV*.
- Corley, M., Stewart, O.W. 2008. Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass*, 2(4), 589-602.
- Eklund, R. 2004. Disfluency in Swedish human-human and human-machine travel booking dialogues. Ph.D. dissertation, Linköping University Electronic Press.
- Lickley, R.J. 2015. Fluency and Disfluency. In Redford, M.A. (Ed.). *The handbook of speech production*, 445-474. John Wiley & Sons.
- Origlia, A., Savy, R., Poggi, I., Cutugno, F., Alfano, I., D’Errico, F., Vincze, L., Cataldo, V. 2018. An Audiovisual Corpus of Guided Tours in Cultural Sites: Data Collection Protocols in the CHROME Project. Proc. 2018 AVI-CH Workshop, 1-4, Castiglione della Pescaia, Italy.

# The natural environment for the experimental study of phonosemantics

Elena Shamina

Department of Phonetics, Saint Petersburg State University, Russia

DOI:

## Abstract

It is claimed that the most natural environment for experimental phonosemantic studies is that of the minimal text in the respondents' native language with the use of real (native or foreign) lexis in its sound form. The claim is based on the idea that if language is viewed as the tool for cognition and communication, then its units, including iconic ones, realize their functions in all their entirety exclusively in a verbal (originally oral) context, while iconic non-language vocalizations or pseudowords serve only as evidence of such psycho-physiological phenomena as synesthesia and synesthesia. The results of the perceptual experiment with English and Russian speakers seem to point to the validity of the claim. Key words: phonosemantics, natural language units, linguistic context.

Key words:

## Background: phonosemantics

The sound form of iconic language signs, sound imitative or sound symbolic ones, (unlike indices or symbols in C.S.Peirce's classification (1994)) represents at least some, though definitely not all, components of their meaning. As their functioning is based on universal cognitive principles (Fodor 1975) and psychological and neuro-physiological foundations, such as synesthesia (Горелов 2003) and synesthesia (Воронин 1982) they may be supposed to have cognitive potential and to provoke correct perception and understanding of foreign words.

## Background: experimental studies

Experimental studies of phonosemantics that have a long and notable history dating back almost a century (cf. (Köhler 1929)) may be said to have been mixing language and non-language environment for the procedure. Very few of them have relied on perception and evaluation of the linguistic icons in context. The methods have varied, including not only simple matching or translation, but also highly standardized numerical ones (Osgood et al 1967). Different types of material have been used: separate speech sounds, graphemes (Журавлев 1974), pseudo-words (Köhler 1929; Горелов 2003), non-language vocalizations (Perlman, Lupyan 2018) and lexis of existing natural languages (Osgood et al 1967). Representation of meaning has been

offered to the subjects with the help of pictures, abstract figures, descriptions and words of the subjects' native tongues or foreign languages. It is worthy of notice that none of these is a regular feature of natural language communication. That is probably why there are instances when the assumed general phonosemantic correspondences between sound and meaning do not come true in a particular language (Syles, Gawne 2017).

Synesthesia and synesthesia may be part of the human psychological and neuro-physiological mechanisms, but the question still remains if they play a significant role in communication in a living language. If language should be viewed as the tool for cognition and communication, then its units, including iconic ones, realize their functions in all their entirety exclusively in a verbal context, presumably in the oral form. This observation sets certain restrictions on the format of the phonosemantic experiments to create the most natural environment for the study.

## **Perception of foreign words in a native language context**

### **Experiment design**

The experiments were specifically designed to ensure that the subjects dealt with *linguistic* instances in their *sound* form, and not just any acoustic or visual stimuli.

They were exposed to utterances in their native tongue that they could see in writing and asked to fill in the gaps by picking one out of the two foreign words presented to them in the (pre-recorded) audio form, e.g., A. *Leaves began to ... (1 свистеть / 2 шелестеть) ... and a door ... (1 звенеть / 2 грохотать) ... shut*; B. *Все девушки обсуждали, как закадрить самых ... (1 гитр / 2 зинг) ... мальчиков на курсе*. The respondents were speakers of Russian and English, of both sexes and of various social and age groups, 82 and 103 people respectively, who did not know the language of the target (sounded) stimuli. The Russian and English sentences were practically identical semantically and contained iconic words that were either sound imitative (example A) or sound symbolic with pejorative or meliorative connotation (example B). These were extracted from the English iconic vocabulary dictionary (Флакман 2016) or the author's corpus of Russian and American expressive slang terms. More than 4300 responses were elicited and analysed and the data statistically evaluated.

### **Results and discussion**

The high percentage of correct choices made by the subjects in the experiments, as well as no cases of refusal to take part in them, definitely point to a cognitive potential of iconic lexis or, in other words, to its euristic function (Шамина 2018). It should also be noted that any observed

discrepancies between reactions of respondents in different age and gender groups are statistically insignificant.

The average correct identification and interpretation of the foreign language units under consideration varies from 70% for Russian sound imitations to 77% for English emotive slang words, peaking to 96% for certain items.

Thus, in respect to onomatopes, English speaking subjects in 87% of cases chose *брякнуть* /br'aknut'/ over *жуужжать* /жуужжать'/ to designate the sound of keys put into a pocket, and Russian speaking subjects preferred the English word *rumble* /rʌmbl/ to *stutter* /stʌtʃr/ in the context of ... *my stomach ... in anticipation of dinner* in 91% of cases. Different classes of onomatopes (instants, continuants and frequentatives – cf. (Voronin 2005)) seem to demonstrate certain tendencies in being recognized by speakers of other languages easier or with more difficulty. Instants, or imitations of pulses, comprise a group of words that are better recognized in both of the languages under study.

Speaking of sound symbolic items, pejoratives and melioratives demonstrate clear-cut phonetic forms that are perceived as such by speakers of the other language. For example, 96 % of the Russian subjects preferred *nerd* /nɜrd/ to *sunny* /sʌnʌ/ as a pejorative, and English subjects did the same in respect to *пентюх* /p'ent'ux/ in contrast to *дельный* /d'el'nɔj/ in 92 % of cases. For some reason, pejoratives turned out to be easier recognized than melioratives.

The results in most cases can be accounted for by the well-established tendencies in language sounds representing non-linguistic acoustic phenomena (e.g. trills imitating vibrations) and non-acoustic notions (e.g. back rounded vowels standing for “bad” things, or labials and velars being “sound gestures” of dislike) that are supposed to be universal (Шамина 2013).

## Conclusion

The data gained from the experiments described above can be interpreted in two ways. First, the results clearly demonstrate the possibility of a nonarbitrary link between the sound form of a (expressive) word and its meaning. Second, the specifically linguistic design of the experiments helps to avoid the influence of purely visual and non-linguistic acoustic factors on the outcome. This makes the evidence for non-arbitrary links between the sound form and the meaning of natural language units elicited in this way more valid linguistically.

## Acknowledgements

The author is deeply grateful to Natalia Dmitrieva for her help in conducting the experiment.

## References

- Fodor, J.A. 1975. *The Language of Thought*. N.Y., Harvard University Press.
- Köhler, W. 1929. *Gestalt Psychology*. New York, Liveright.
- Peirce, C. S. 1994. In *Peirce on Signs: Writings on Semiotic*. James Hoopes (ed.) 1994, Chapel Hill, NC, University of North Carolina Press.
- Perlman, M., Lupyán, G. 2018. People can create iconic vocalizations to communicate various meanings to native listeners. *Scientific Reports* 8 (1), 2634.
- Syles, S., Gawne, L. 2017. When Does Maluma/Takete Fail? Two Key Failures and a Meta-Analysis Suggest That Phonology and Phonotactics Matter. *I-Perception* 8 (4), 2041669517724807.
- Osgood, C.T., Suci, G.J., Tannenbaum, P.H. 1967. *The Measurement of Meaning*. Urbana, University of Illinois Press.
- Voronin, S.V. 2005. The Universal Classification of Onomatopoes 25 Years On. In *Iconicity. Glottogenesis. Semiosis*. St.Petersburg, St.Petersburg State University Press. 30 – 37.
- Воронин С.В. 1982. Основы фоносемантики. Ленинград, Изд-во ЛГУ. Горелов И.Н. 2003. Проблема функционального базиса речи в онтогенезе. In *Избранные труды по психолингвистике*. Москва, Лабиринт, 15-104.
- Журавлев А.П. 1974. *Фонетическое значение*. Ленинград, Изд-во ЛГУ.
- Флакман М. А. 2016. *Словарь английской звукоизобразительной лексики в диахроническом освещении*. СПб., Издательство РХГА.
- Шамина Е.А. 2013. Восприятие эмоционально-оценочной лексики неродного языка: универсальные и частноязыковые факторы. In *Proc. 10<sup>th</sup> International Congress of the International Society of Applied Psycholinguistics*. 272-273. Moscow, Russia.
- Шамина Е.А. 2018. Когнитивная валентность звукоподражательной лексики (экспериментальное исследование). In *Когнитивные исследования языка, XXXIY. Cognito и Communicatio в современном глобальном мире*. Москва – Тамбов, Ин-т языкознания РАН, 338-342.

# Analysing L2 Swedish word-final stops

Maryann Tan<sup>1</sup>, Xin Xie<sup>2</sup>, T. Florian Jaeger<sup>2</sup>

<sup>1</sup>Centre for Research on Bilingualism, University of Stockholm, Sweden

<sup>2</sup>Brain & Cognitive Sciences, University of Rochester, USA

DOI:

## Abstract

We compare native (L1) and non-native (L2) word-final plosive voicing in Swedish. The L1 of the L2 speaker (Flemish) does not have word-final plosive voicing contrasts. In order to assess the effectiveness of a common approach to L2 instruction, L2 speech was elicited under two conditions: either unassisted or by playing an example L1 production and asking the L2 speaker to mimic it. Three cues to voicing—vowel, closure, and burst durations—were measured. L2 productions relied on different cues for voicing than L1 production. Mimicking reduced the difference between L1 and L2 speech.

Key words: speech, L2, Swedish, mimicking, word-final stop voicing

## Introduction

Acquiring native-like pronunciation is one of the most difficult feats for L2 learners. Non-native accents are pervasive even after years of L2 learning, with potential social consequences (Fuertes et al., 2009). Formal instruction is not necessarily a significant predictor of accentedness, partly because pronunciation is not prioritised in foreign language teaching (Piske et al., 2001).

In a single-talker case study, we assess the effectiveness of a common instructional approach to L2 accent reduction—repeating (‘mimicking’) words immediately after hearing the native pronunciations. The data were generated as stimuli for an ongoing project on native perception of L2 accented speech. Specifically, we compare native and non-native realizations of word-final stop voicing in Swedish, when the L1 of the L2 speaker (Flemish) has no such voicing contrast.

## Materials and method

Speech stimuli comprised 32 minimal pairs of monosyllabic Swedish words (e.g. röt-röd), differing only in final stop voicing. As vowel length is phonemic in Swedish, the words of a minimal pair were always matched in phonological vowel length.

Recordings were made of a 35-year old, female native speaker. The speaker was from the north of Sweden but had lived in Stockholm for 10 years prior to the recordings. Her accent was judged as close to Central

Standard Swedish (CS) by native speakers of CS. The L2 speaker was a 25-year old female native speaker of the Central Flandrian dialect of North Belgium, with only level A1 (CEFR) knowledge of Swedish.

Recordings were made in a sound-attenuated room at the Stockholm University Multilingualism Lab. Written forms of the words were presented on a computer screen placed within a comfortable viewing distance from the speaker. Words were spoken into an Audio-Technica AT3035 microphone, placed directly in front of the speaker. Recordings were sampled at 44.1kHz. The experimenter controlled the presentation of each word.

In sessions with the native speaker, the 64 words were presented twice in random order. The speaker was prompted to produce the word when it was displayed onscreen. The two recording conditions (unassisted vs. mimicked) with the L2 speaker are described next. In both conditions, voiced and voiceless words were recorded separately in order to ease the production task for the L2 speaker.

### **Unassisted speech**

The procedure was identical to the recordings of the native speaker, except that the 64 target words were presented three times in randomised order. This decision was made because the speaker was expected to produce a higher rate of disfluent productions, which cannot be reliably annotated.

### **Mimicking**

Recordings were made in a session one week after the unassisted recordings. The procedure was identical to that used in the unassisted condition, with one exception: at the beginning of each trial, a recording of the native speaker producing the target word was played to the L2 speaker over Sony MDR-7506 headphones at a comfortable volume. Simultaneously and throughout the trial, the target word was displayed onscreen. An audible beep was played after 2 seconds from trial onset (after the native recording had finished playing) to cue production of the target.

### **Annotation**

Here we analyse the first 64 tokens from each of the three conditions (L1, mimicked L2, unassisted L2). Tokens were annotated for their duration of vowel, closure, and burst. Annotations were completed in Praat (Boersma, 2001) using visual examination of spectrograms, and listening judgments. Cue boundaries were marked following conventions (Flege, Munro, & Skelton, 1992). Vowel duration was measured from the beginning of the first periodic portion of each waveform to the zero-crossing where the amplitude decreased abruptly and the waveform became sinusoidal. Burst was measured from stop release to the first zero crossing point where the

amplitude became near zero. Closure was measured as the time between vowel offset and burst onset (for stops following nasals, closure onset was marked by an abrupt decline in amplitude of the nasal). Four tokens in the unassisted condition had mispronounced vowels. These were replaced by one of the other two recordings of the same word (recall that we recorded each token three times for the non-native speaker).

## Results

We compared unassisted against native speech, and against mimicked speech. Separate analyses were conducted for the three measures. All analyses used linear mixed-effects regression with maximal random effect structures (random intercepts and slopes for elicitation condition by minimal pair). Voicing was sum-coded (/d/=1 vs. /t/= -1) and elicitation treatment-coded (baseline: unassisted condition). We focus on interactions of voicing and elicitation, as this captures differences in the realization of voicing.

### Native vs unassisted L2 speech

All three measures exhibited a significant interaction ( $ps < .006$ ). Critically, the pattern of the interaction differed across the three measures. Simple effect analysis found that vowels were shorter for /d/- than for /t/-final words only in L2 speech ( $p < .03$ ), but not L1 speech ( $p > .11$ ). Both closure and burst were significantly shorter for /d/ than /t/ both for unassisted and mimicked L2 speech ( $ps < .006$ ), but this difference between /d/ and /t/ was 3-times larger for native speech for closure, and 3-times larger for *unassisted* speech for bursts. The L2 speaker is thus non-native both in terms of what cues she relied on, and how they are used. Specifically, she primarily relied on a cue secondary to native speakers (burst), instead of the primary native cue (closure).

### Unassisted L2 speech vs. mimicked L2 speech

All measures exhibited a significant interaction ( $ps < .0002$ ). Simple effect analysis found that vowels were longer for /d/ than /t/ in the mimicked condition ( $p < .0001$ ), but shorter in the unassisted condition ( $p < .03$ ). Both closure and burst were significantly shorter for /d/ than /t/ both for unassisted and mimicked speech ( $ps < .001$ ). But this effect was larger for mimicked speech for closure, and larger for unassisted speech for bursts. Overall, mimicked speech more closely resembled native speech (separate analyses found that mimicked speech still significantly differed from L1 speech, except for closure durations).

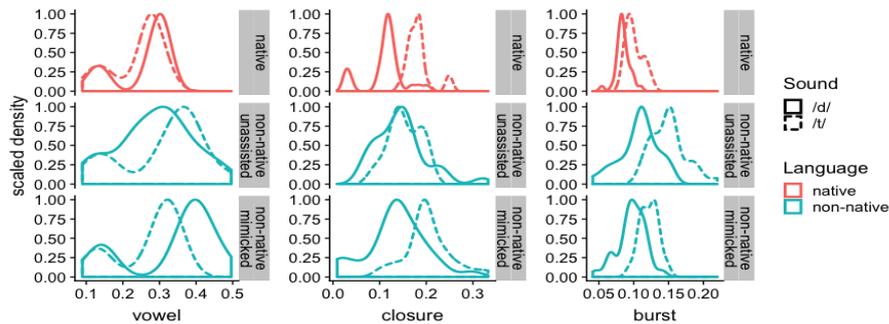


Figure 1. Density of cue durations in voiced and voiceless word pairs for native (top row), unassisted non-native (middle) and mimicked non-native speech (bottom).

## Discussion and conclusion

This presents the first (small-scale) quantitative analysis of L2 word-final stops in Swedish (previous work has used phonological judgments, Hammarberg, 1997). We find that immediate repetition substantially reduces the non-nativeness of pronunciations, though our results leave open how long lasting this effect is. The specific non-native pattern we find for L2 Swedish pronunciations closely resembles the pattern found for L2 English final stop voicing by L1 Mandarin speakers (Xie et al., 2017). Both Flemish and Mandarin have stop voicing only in initial positions. The similarity in the L2 speech across these two studies might suggest that Flemish and Mandarin L2 learners transfer L1 knowledge about initial stop voicing (which relies on burst/voice onset time) to the production of L2 final stops.

## References

- Fuertes, J.N., Gottdiener, W.H., Martin, H., Gilbert, T.C., Giles, H. 2012. A meta-analysis of the effects of speakers' accent on interpersonal evaluations. *Eur. J. Soc. Psychol.* 42, 12–133.
- Boersma, P., 2001. Praat, a system for doing phonetics by computer. *Glott International*, 5 (9/10) (2001)
- Flege, J., Munro, M., Skelton, L. 1992. Production of the word-final /t/-/d/ contrast by native speakers of English, Mandarin, and Spanish. *JASA*, 92, 128-143.
- Hammarberg, B. 1997. Conditions on transfer in phonology. In A. James and J. Leather (Eds.), *Second-language speech* (161–180). Berlin: Mouton de Gruyter.
- Piske, T., MacKay, I.R.A., Flege, J.E. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29 (2), 191–215.
- Xie, X., Theodore, R.M., Myers, E.B. 2017. More than a boundary shift. *JEP:HPP*, 43(1), 206–217.

# Speech and Mild Cognitive Impairment detection

Charalambos Themistocleous<sup>1,2</sup>, Dimitrios Kokkinakis<sup>2,3</sup>

<sup>1</sup> Department of Neurology, Johns Hopkins University, USA

<sup>2</sup> Department of Swedish, University of Gothenburg, Sweden

<sup>3</sup> Centre for Ageing and Health, University of Gothenburg, Sweden

DOI:

## Abstract

It is of great importance to detect objective markers that can enable the early and fast identification of individuals with Mild Cognitive Impairment (MCI) from healthy individuals to inform, patient care, family and treatment planning. Connected speech productions can offer such markers. This study analyses recordings from picture description tasks by Swedish individuals with MCI and healthy control individuals (HC) and shows that voice quality, periodicity, and speech rate distinguish individuals with MCI from HC.

Key words: Mild Cognitive Impairment, Alzheimer's disease, speech analysis

## Introduction

Mild Cognitive Impairment (MCI) does not always lead to Alzheimer's Disease (AD) but in most cases it does (Mitchell, Shiri-Feshki 2009). The challenge is to discover subtle markers early on that can identify individuals with MCI. A number of studies show that speech and language can provide such early markers (Asgari et al. 2017, Dodge et al. 2015, Roark et al. 2011, Satt, et al. 2014). Most studies focused primarily on acoustic markers that distinguish linguistic categories (for other features see e.g., Fraser et al. 2016), such as vowel formant frequencies, segmental duration (long/short vowels), and intonation and showed that they can provide early cues for MCI. In this line, we have developed a Deep Neural Network that identifies individuals with MCI and healthy controls (HC) with high classification accuracy ( $M = 83\%$ ) by using only a couple of acoustic features (i.e., vowel duration, formant frequencies, and F0) (Themistocleous et al. 2018). Nevertheless, more research is required on acoustic features that associate with non-linguistic properties. The aim of this study is to identify early markers of MCI from voice quality, speech rate, and periodicity.

## Methodology

Thirty-two individuals with MCI and third-two HC individuals between 55 and 79 years old ( $M = 69$ ,  $SD = 6.4$ ) participated in the study; these individuals did not differ with respect to age [ $t(52.72) = -1.8178$ ,  $p = n.s.$ ] and gender ( $W = 1567.5$ ,  $p = n.s.$ ). Six selection criteria were followed: (i)

participants should be Swedish speakers; (ii) be able to read understand information about the study; (iii) provide written consent and not suffer from (iv) dyslexia and reading difficulties; (v) major depression, ongoing substance abuse; (vi) other psychiatric, neurological or brain-related conditions. Healthy control individuals had higher Mini-Mental State Exam (MMSE) score. Ethic approvals for the study were obtained by the local ethical committee review board (reference number: L091-99, 1999; T479-11, 2011); while the currently described study was approved by the local ethical committee decision 206-16, 2016. The recordings were conducted in two phases the first phase begun in late 2016 and the second in 2018. We have measured (i) the overall production of MCI vs. HC (“Speaking Time”); (ii) the articulation rate of these two groups; (iii) voice quality measures, namely the difference between the first harmonic and third amplitude (H1-A3); and measures of periodicity, namely the Cepstral Peak Prominence (CPP). Higher values of CPP correspond to greater periodicity. We report linear regression models. In the tables, (1) significant  $b$ -weights indicate, a significant semi-partial correlation. (2)  $b$  stands for the unstandardized regression weights; (3)  $sr^2$  represents the semi-partial correlation squared; (4)  $LL$  and  $UL$  indicate the lower limit of confidence intervals and their upper limits. One star \* indicates  $p < .05$ ; two stars \*\* indicate  $p < .01$ .

Table 1. Regression results for speaking time.

Predictor	$b$	$b$ 95% CI [LL, UL]	$sr^2$	$sr^2$ 95% CI [LL, UL]	Fit
(Intercept)	38.28**	[32.86, 43.71]			
conditionMCI	11.38**	[3.19, 19.57]	.06	[.00, .16]	
					$R^2 = .063^{**}$ 95% CI [.00, .16]

## Results

MCI individuals produced slower speech; the effect resulted in statistically significant differences of individuals with MCI from HC (see Table 1).

Table 2. Regression results for articulation rate.

Predictor	$b$	$b$ 95% CI [LL, UL]	$sr^2$	$sr^2$ 95% CI [LL, UL]	Fit
(Intercept)	65.47**	[58.46, 72.48]			
conditionMCI	-14.29**	[-24.87, -3.71]	.06	[.00, .16]	
					$R^2 = .060^{**}$ 95% CI [.00, .16]

Individuals with MCI had slower articulation rate than HC individuals (Table 2). Also, there were significant effects on voice quality effects. The H1-A3 was significantly different for individual factors of condition and gender and their interaction (Table 3). The results on CPP were significant for condition, gender, and their interaction. Especially, there was greater CPP for female HC speakers than female speakers with MCI (see Table 4).

Table 3 Regression results for H1-A3 (gender M=Male).

Predictor	<i>b</i>	<i>b</i> 95% CI [LL, UL]	<i>sr</i> <sup>2</sup>	<i>sr</i> <sup>2</sup> 95% CI [LL, UL]	Fit
(Intercept)	26.87**	[25.14, 28.59]			
conditionMCI	2.91*	[0.52, 5.29]	.02	[-.01, .05]	
genderM	9.75**	[6.11, 13.39]	.09	[.03, .15]	
conditionMCI: genderM	-5.52*	[-10.06, -0.97]	.02	[-.01, .05]	
					<i>R</i> <sup>2</sup> = .137**
					95% CI [.06, .21]

Table 4 Regression results for CPP (gender M=Male).

Predictor	<i>b</i>	<i>b</i> 95% CI [LL, UL]	<i>sr</i> <sup>2</sup>	<i>sr</i> <sup>2</sup> 95% CI [LL, UL]	Fit
(Intercept)	71.12**	[70.59, 71.66]			
conditionMCI	-1.18**	[-1.92, -0.44]	.03	[-.01, .07]	
genderM	-1.96**	[-3.08, -0.83]	.04	[-.00, .08]	
conditionMCI: genderM	2.05**	[0.64, 3.46]	.03	[-.01, .07]	
					<i>R</i> <sup>2</sup> = .057**
					95% CI [.01, .11]

## Discussion

This study analysed acoustic productions from recordings of a picture description task. Overall, individuals with MCI produce speech which is slower than the speech of HC individuals. Earlier studies suggest that MCI individuals produce longer pauses and longer segmental duration (see also Slegers et al. 2018). Our finding may suggest that longer vowel duration and longer pause duration can be attributed to the overall slower speech. Individuals with MCI differ from HCs on voice quality as well. These differences were shown on the difference between the first harmonic and third amplitude that associates with phonation Esposito (2010). Finally, the periodicity of speech in individuals with MCI and HC was found to differ with respect to CPP. Especially, there was lower CPP in female individuals with MCI than in female and male HC individuals, but the effects were less

clear with respect to male individuals. Acoustic markers such as vowel formant frequencies that distinguish vowel quality, the fundamental frequency, which is the main acoustic correlate of intonation, segmental duration, etc. were studied extensively. These findings show that non-linguistic markers of speech production, such as speech periodicity, speech rate, and measures of voice quality can be employed to identify individuals with MCI from HC individuals. In our future research, we plan to integrate linguistic and non-linguistic acoustic properties elicited automatically from speech signals to identify individuals with MCI from HC individuals.

### **Acknowledgements**

This research has been funded by Riksbankens Jubileumsfond – The Swedish Foundation for Humanities and Social Sciences (NHS 14-1761:1).

### **References**

- Asgari, M., Kaye, J., Dodge, H. 2017. Predicting mild cognitive impairment from spontaneous spoken utterances. *Alzheimer's and Dementia: Translational Research & Clinical Interventions*, 3, 219–228.
- Dodge, H., Mattek, N., Gregor, M., Bowman, M., Seelye, A., Ybarra, O., Kaye, J. 2015. Social markers of mild cognitive impairment. *Current Alzheimer Research*, 12, 513–519.
- Esposito, C.M. 2010. The effects of linguistic experience on the perception of phonation. *Journal of Phonetics* 38, 306–316.
- Fraser K.C., Meltzer J.A., Graham N.L., Leonard C., Hirst G., Black S.E., Rochon E. 2014. Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex* 55, 43-60.
- Kokkinakis, D., Lundholm Fors, K., Fraser, K., Eckerström, M., Horn, G., Themistocleous, C. 2019. A multifaceted corpus for the study of cognitive decline in a Swedish population. In CLARE4. Helsinki, Finland.
- Mitchell, A.J., Shiri-Feshki, M. 2009. Rate of progression of mild cognitive impairment to dementia – meta-analysis of 41 robust inception cohort studies. *Acta Psychiatrica Scandinavica*, 119: 252-265.
- Roark, B., Mitchell, M., Hosom, J.P., Hollingshead, K., Kaye, J. 2011. Spoken language derived measures for detecting mild cognitive impairment. *IEEE transactions on audio, speech, and language processing*, 19(7), 2081-2090.
- Satt, A., Hoory, R., König, A., Aalten, P., Robert, P.H. 2014. Speech-based automatic and robust detection of very early dementia. In Fifteenth Annual Conference of the International Speech Communication Association.
- Slegers, A., Filiou, R.P., Montembeault, M., Brambati, S. M. 2018. Connected Speech Features from Picture Description in Alzheimer's Disease: A Systematic Review. *Journal of Alzheimer's Disease* 65: 519–542.
- Themistocleous C., Eckerström M., Kokkinakis, D. 2018. Identification of Mild Cognitive Impairment from Speech in Swedish using Deep Sequential Neural Networks. *Frontiers in Neurology* 9: 975.

# Hungarian demonstratives in contrastive contexts

Enikő Tóth

Department of English Linguistics, University of Debrecen, Hungary

DOI:

## Abstract

This paper presents the results of corpus queries and a subsequent questionnaire study exploring the exophoric use of four Hungarian demonstratives *ez/az* ‘this/that’ and *emez/amaz* ‘this/that other one’ in contrastive contexts. Concordance analysis revealed that the aforementioned four demonstratives often occur in various patterns, such as *ez-amaz*. Findings of the questionnaire study showed that in contrastive contexts, where two objects are placed either within or out of arm’s reach in front of the speaker and addressee in table-top space on the sagittal axis, the acceptability of the emerging contrastive patterns does not depend on relative distance. This means that speaker proximity can be overridden by contrastive function in Hungarian.

Key words: demonstratives, contrastive, exophoric, Hungarian

## Introduction

Hungarian has a two-way speaker-proximal versus distal system of demonstratives, comprising proximal *ez* ‘this’ and distal *az* ‘that’. However, there are two other, less frequently used demonstratives, proximal *emez* ‘this one nearby/this other one’ and distal *amaz* ‘that one over there/that other one’, which are either assumed to make a more finer distinction within near and far, or to have a reinforcing role (Laczkó 2012). The aim of this paper is to explore a novel, previously neglected contrastive use of *emez/amaz*, relying on corpus based findings and experimental data.

Crosslinguistically, exophoric contrastive uses of demonstratives, such as *I find this sofa nice, but that one is ugly*. (cf. Levinson 2004), have been studied for example by Meiro and Terrill (2005), who showed that the use of demonstratives is different in contrastive and non-contrastive contexts in Tiriyo and Lavukaleve. Further elicitation studies revealed that in contrastive contexts proximity is often neutralized across languages (see Wilkins 1999 and Levinson 2018 for details). In the case of Hungarian, Tóth et al. (2014) compared the use of the more frequent demonstratives (*ez/az*) in neutral vs. contrastive contexts (*Ezt kérem*. ‘I want this.’ vs. *Ezt a dobozt vidd le a pincébe, azt viszont hagyd a helyén*. ‘Take this box down to the cellar, but leave that one where it is.’ Laczkó 2012: 296). It was shown that while relative distance from the speaker is crucial in neutral contexts, in contrastive contexts, when the objects being referred to are both close to the speaker, the distal term (*az*) is used in a significantly higher proportion. The

aim of this study is to widen the scope of investigation and to examine the use of *emez/amaz* in contrastive contexts.

### Corpus findings

First, frequency queries were run in the Hungarian National Corpus, which found 728 and 2331 tokens of *emez* and *amaz*, respectively. Concordance analysis revealed that each of the four demonstratives can create a contrast with other referring expressions. There were altogether 943 contrastive examples including *emez* or *amaz* in the corpus. *emez/amaz* typically appeared in contrastive contexts in patterns where one proximal and one distal term are combined. A relevant example is the following: *Ez ide az ágy alá, amaz meg a polcra – jó lesz karácsonyra*. ‘This comes here below the bed, that other one goes to the shelf – for Christmas’. *ez-emez*, *az-amaz* might also surface in the same utterance, but these patterns were not so frequent. These findings are in line with cross-linguistic data (cf. Wilkins 1999, Maes & de Rooij 2007). Table 1 shows the relative frequency distribution of the patterns emerging.

Table 1. Relative frequency of patterns including *emez/amaz* in HNC.

pattern	number	relative frequency
amaz-amaz	1	0.1 %
az-emez	35	3.7 %
az-amaz	227	24 %
emez-amaz	141	15 %
emez-emez	16	1.7 %
ez-emez	27	2.9 %
ez-amaz	496	52.6 %

### Questionnaire study

The most frequent five patterns in Table 1 served as input to an online questionnaire study. The question to be addressed was whether the acceptability of the patterns depends on the relative distance of the objects being referred to from the speaker, i.e. whether the contrastive function of the patterns can overwrite the role of proximity. 55 subjects saw photos depicting the position of two speakers (sitting next to each other) with respect to two objects placed in table-top space on the sagittal/away axis. One of the referents was object-identifiable, the other was referred to by a demonstrative pronoun. Participants provided acceptability judgements of contrastive utterances exemplifying one of the patterns on a 5 point Likert scale. Each target utterance contained two demonstratives (i.e. one pattern

was presented), such as *Az a fekete esernyő sokkal strapabíróbb modell, mint emez*. ‘That black umbrella is sturdier than this other one.’ The referents were easily identifiable, and there was no difference between the objects regarding salience.

The questionnaire contained 20 target items and 10 fillers. Two factors were examined in a 2x5 design: DISTANCE: peripersonal (within arm’s reach) vs. extrapersonal (out of arm’s reach) (see Kemmerer 1999), and TYPE OF CONSTRUCTION: *ez-emez*, *ez-amaz*, *az-amaz*, *az-emez* and *emez-amaz*. Overall results are presented in Figure 1.

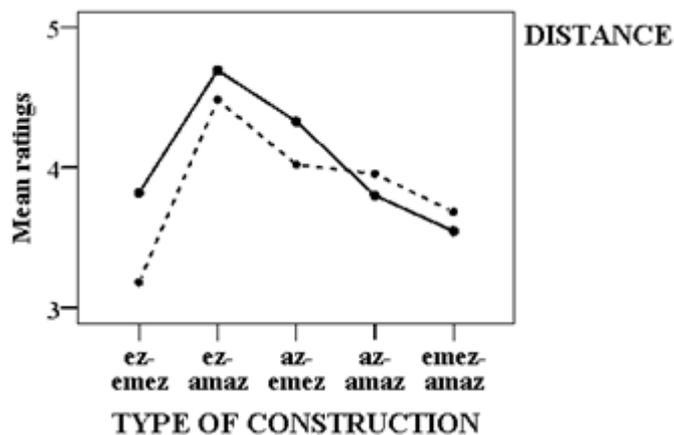


Figure 1. Overall results.

*ez-amaz* received the highest average ratings in both distance conditions and this pattern was also the most frequent one in the Hungarian National Corpus. There was only a marginal effect of DISTANCE, this factor is only responsible for 15 % of overall variability:  $F(1, 54) = 9.519$ ,  $p < 0.05$ ,  $\eta^2 = 0.15$ . There were significant differences in between the peripersonal and extrapersonal conditions only in the case of *ez-emez*, and *az-emez*, both received significantly higher ratings in the peripersonal condition. Regarding *ez-emez* this is not surprising, since a combination of two proximals is preferred when referring to entities that are within arm’s reach from the speaker. Considering *az-emez*, *az*, a distal demonstrative creates a contrast with an entity that is closer to the speaker within peripersonal space.

In the case of the remaining patterns, distance is not a decisive factor, i.e. in contrastive contexts distance as a factor can be overridden; the patterns in question are equally acceptable both in the peripersonal and extrapersonal conditions. The findings reinforce the results of the previous study (Tóth et al. 2014) on the contrastive uses of *ez/az*, and are in line with cross-linguistic results. Namely, it has been shown that in contrastive contexts proximal

demonstratives are also acceptable when they refer to objects that are far from the speaker, while distal terms are acceptable when the referents are within arm's reach.

A more substantial main effect of TYPE OF CONSTRUCTION was also found,  $F(1, 51) = 22.533$ ,  $p < 0.001$ ,  $\eta^2 = 0.64$ . *ez-amaz*, the most frequent pattern in the corpus study, received significantly higher ratings than each of the other patterns, while the least frequent pattern, *ez-emez*, got the lowest ratings. There was also a significant interaction between the two variables.

To conclude, the results of the two studies report similar findings. Each of the Hungarian demonstratives (*ez/az*, *emez/amaz*) appears in contrastive contexts, and the findings of the rating task revealed that speaker proximity can be overridden by contrastive function in Hungarian.

### Acknowledgements

The Project no. 111918 (*New approaches in the description of the grammar of Hungarian pronominals*) has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary.

### References

- Kemmerer, D. 1999. 'Near' and 'far' in Language and Perception. *Cognition* 73, 35-63.
- Laczkó, K. 2012. Spatial deixis and demonstrative pronouns in Hungarian. In Hart, C. (ed.) *Selected Papers from UK-CLA Meetings*, 289-301, <http://uk-cla.org.uk/proceedings>
- Levinson, S. C. 2004. Deixis. In Horn, R., Ward, R. (eds.) *The Handbook of Pragmatics*, 97-121, Oxford, Blackwell Publishing.
- Levinson, S. C. 2018. Introduction: Demonstratives: Patterns in Diversity. In Levinson, S. C., Cutfield, S., Dunn, M. J., Enfield, N. J., Meira, S. (eds.) *Demonstratives in Cross-Linguistic Perspective*, 1-42. CUP.
- Maes, A., de Rooij, C. 2007. (How) Do demonstratives code distance? In Branco, A., McEnery, T., Mitkov, R., Silva, F. (eds.) *Proc. of DAARC 2007*, 83-89, Lagos Portugal, Centro de Linguística da Universidade de Porto.
- Meira, S., Terril, A. 2005. Contrasting contrastive demonstratives in Tiriyó and Lavukaleve. *Linguistics* 43, 1131-1152.
- Tóth, E., Csatár, P., Banga, A. 2014. Exploring Hungarian and Dutch gestural demonstratives. In Veselovská, L., Janebová, M. (eds.) *Complex Visibles Out There. Proc. of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure*. 607-625, Olomouc: Palacký University.
- Wilkins, D. 1999. Eliciting contrastive use of demonstratives for objects within close personal space (all objects well within arm's reach). In Wilkins, D. (ed.) *Manual for the 1999 Field Season*. 25-28, Nijmegen: Max Planck Institute for Psycholinguistics.

# Lexical selection and gender agreement processing

Masayuki Yamada, Maria do Carmo Lourenço-Gomes  
Institute of Arts and Humanities, University of Minho, Portugal  
DOI:

## Abstract

This study investigates the influence of lexical selection of adjectives on gender agreement in writing by late Chinese learners of European Portuguese (EP), with an online paradigm. Participants were asked to translate sentences (from Chinese to Portuguese) containing adjectives that ranged from easy to difficult, on a platform that recorded the writing process. A greater pause was observed in the *difficult* condition than in the *easy* condition at the critical position, i.e., just before the adjective. In addition, there were more errors in the *difficult* condition. These results suggest that when lexical selection is difficult it demands more resources from working memory (WM), and, hence, less resources are left for the agreement process, which results in more errors.

Keywords: gender agreement, lexical selection, Portuguese, Chinese

## Background

Acquisition of gender agreement by late L2 learners has generated much research in the literature. There are, however, few studies on gender agreement in EP. Most have examined written corpora data and discussed the relationship between errors and noun features such as morphology and animacy (e.g., Brito, 2015; Pinto, 2014). Despite their relevant findings, their results should be interpreted with caution, because they did not distinguish between *attribution* errors and *agreement* errors in the analysis, which renders difficult to discuss the nature of the errors (whether lexical or syntactic). Furthermore, as these analyses were conducted on the final version of the texts, the online process of writing was not examined. For these reasons, Yamada et al. (2018) analysed gender agreement errors in writing by late Asian learners of EP, by using a platform developed by the author (demo: <http://bit.ly/2M4FCgR>), which enabled the recording and later review of how the writing process unfolded in real time due to a keystroke logging technique (for a review, see Van Waes, Leijten, Wengelin, & Lindgren, 2011). As for noun-adjective agreement, the results suggested an independent relationship between noun's features (transparency and animacy) and occurrence of *agreement* errors. Moreover, an analysis of the writing process revealed that a relatively great pause was found just before the occurrence of some errors. A post-interview to the participants unveiled

that they were engaged in the lexical selection of adjectives during the pause, which suggests an influence of adjectives on gender marking. To our knowledge, however, the influence of adjective features on noun-adjective agreement in production has not been studied yet.

As far as agreement errors are concerned, the role of WM may be considered as well. Fayol et al. (1994) found that, when a cognitive load was imposed by a concurrent task, native French speakers made more errors of subject-verb agreement. The authors argued that the WM overloaded because of the secondary task, which made correct agreement marking difficult. McDonald (2006) reported that decreased performance of L1 speakers under cognitive load mirrors that of L2 speakers, positing that poor performance of L2 speakers resulted from a short availability of WM's resource. Based on these findings, we postulated that, if the lexical selection of adjectives was causing an overload of learners' WM and consuming its resources to a considerable extent, *agreement* errors were induced by it. This speculation is consistent with evidence showing that lexical process interferes with a later syntactic processing in comprehension (Hopp, 2016).

Considering the above, in the present study, we conducted a translation task (from Chinese to Portuguese) in which the difficulty of translation of adjectives was manipulated to impose cognitive load on lexical selection in order to verify whether such load induced *agreement* errors.

## Study

### Method

Forty-one native speakers of Chinese, late learners of Portuguese as a foreign language, recruited at Portuguese universities, participated in the study (Years of learning: 2.94. Mean-age: 22.6). Participants were asked to translate sentences from Chinese to Portuguese using a writing platform (the same used in Yamada et al., 2018). Target materials for the experiment consisted of six pairs of sentences with the structure: NP + VP (V + AdjP), in which gender agreement between predicate adjective and the sentential subject is obligatory. Each sentence pair differed only in the adjective position. One version of the sentence contained an adjective that could be easily translated (*easy* condition), and the other version contained an adjective that did not translate easily (*difficult* condition). Adjectives were selected from a questionnaire previously conducted with 20 Chinese learners, in which adjectives were rated according to ease/difficulty of translation. An example of a sentence pair is given below (1a-b):

- (1) a. 葡萄牙的生活非常平静 (easy condition)  
b. 葡萄牙的生活非常充实 (difficult condition)  
‘A vida em Portugal é muito calma [easy]/proveitosa [difficult]  
‘Life in Portugal is very calm / profitable.’

A total of 24 sentences (one version of each sentence pair and 18 fillers) was assigned to two lists. Each participant received only one list.

A MacBook Pro (13-inch) with US keyboard was used. To familiarize participants with the task and all the procedures, two training sessions were implemented prior to the main experiment. On each trial each sentence in Chinese appeared at the top of the screen, and participants were asked to translate it into Portuguese in the text area and click the “next” button to access the next sentence. To control for possible fatigue, a break was inserted after half of the sentences. To confirm the appropriateness of the adjectives selected for the task, a post-questionnaire was conducted at the end of the main session, in which participants were asked to rate ease of translation of the target adjectives in a 5-point scale (from 1-very difficult to 5-very easy).

It was hypothesized that a greater pause before the adjective position and more errors (e.g., use of the default form: masculine) would occur in the difficult condition in comparison to the easy condition.

## Results

As expected, a longer pause was observed in the *difficult* condition than in the *easy* condition (Mean-duration: 8738 ms vs. 2153 ms) just before the adjective. Because of high skewness and kurtosis and some outliers, a paired t-test was performed on the log-transformed data, revealing a significant difference between the pause times of the two conditions ( $t(119) = -11.767$ ;  $p < .0001$ ;  $d = 1.53$ ). A Chi-square test revealed that there were significantly more errors in the *difficult* condition than in the *easy* condition ( $\chi^2(1) = 47.053$ ,  $p < .0001$ ,  $\Phi = .447$ ). The mean score of the ease of translation for the easy and the difficult conditions was 4.85 and 2.95 points, respectively. A Wilcoxon signed-rank test revealed that this difference was significant ( $Z = 8.210$ ;  $p < .0001$ ;  $r = .74$ ). A strong correlation between ease of translation and experimental condition was confirmed ( $r = -.715$ ;  $p < .0001$ ).

## Discussion and conclusion

The results obtained indicate that: (i) the choice of adjectives for each condition was appropriate; (ii) the *difficult* condition caused a cognitive load on lexical selection, leading participants to produce longer pauses at the critical position as compared to the easy condition; and (iii) a higher error

rate in the *difficult* condition was observed. Our results seem to indicate that lexical selection may interfere with processing of agreement, which may be taken as evidence for the vulnerability of syntactic processing to processes from other cognitive levels (e.g., Badecker & Kuminiak, 2007). We hypothesize that, as learners' agreement processing is not as automatic as that of native speakers, it may require more resources from WM; when many WM resources are required by other demanding linguistic processes, errors will occur due to the lack of resources available.

Despite some limitations of this study (e.g., reduced number of experimental items), we have replicated experimentally a type of gender agreement error previously observed in spontaneous writing (Yamada et al., 2018). We believe our findings may contribute to the L2 teaching/learning field and to the development of theoretical models of writing.

## Notes

This study is part of the first author's ongoing PhD thesis.

## References

- Badecker, W., Kuminiak, F. 2007. Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language*, 56(1), 65–85.
- Brito, E. 2015. Grammatical gender in the interlanguage of English-speaking learners of Portuguese. *Portuguese Language Journal*, 9, 1–21.
- Fayol, M., Largy, P., Lemaire, P. 1994. When cognitive overload enhances subject–verb agreement errors. *The Quarterly Journal of Experimental Psychology Section A*, 47(2), 437–464.
- Hopp, H. 2016. The timing of lexical and syntactic processes in L2 sentence comprehension. *Applied Linguistics*, 37(5), 1253–1280.
- McDonald, J.L. 2006. Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55(3), 381–401.
- Pinto, J. 2014. A aquisição do género e da concordância de género em português língua terceira ou língua adicional. In P. Osório & F. Bertinetti, *Teorias e Usos Linguísticos*. Lisboa: Lidel.
- Van Waes, L., Leijten, M., Wengelin, Å., Lindgren, E. 2011. Logging tools to study digital writing processes. In V. W. Berninger (Ed.), *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology*. Psychology Press.
- Yamada, M., Lourenço-Gomes, M. do C., Barbosa, P. 2018. Uma análise de tempo real da concordância de género na produção escrita por aprendentes de PLE. Presented at the XXXIV ENAPL 2018, Lisboa, Universidade Aberta.

# Interpreting Mandarin positive polar questions

Chenjie Yuan, Peng Li

Department of Translation and Language Sciences, Universitat Pompeu Fabra, Spain

DOI:

## Abstract

This study examined the non-truth-conditional meaning encoded by sentential-final particles and boundary tones in four Mandarin positive polar question (PPQ) patterns. Two short experiments were conducted to test (a) the role of epistemic and evidential biases (Experiment I) and (b) the antecedent and answer requirements, as well as the (non-)acceptance of the antecedent (Experiment II) in constraining the interpretation of PPQs. In each experiment, a series of situational judgment tasks (27 in Experiment I and 36 in Experiment II) were designed to elicit data from 60 native speakers. The findings suggest that the four patterns are not simple biased questions but are complex non-canonical questions whose interpretation are sensitive to the discourse configuration established by the five abovementioned contextual factors.

Key words: non-truth-conditional meaning, polar question, ending particle, boundary tone, discourse configuration

## Introduction

In the standard alternative semantics of questions (Hamblin 1973; Karttunen 1977), both positive and negative polar questions (PQs) are assigned the same denotation in terms of their alternative answers. Yet, it is widely accepted that both types of PQs are not identical by virtue that PQs may convey information regarding the questioner's bias towards a particular answer and are sensitive to contextual available information (e.g., Ladd 1981; Büring and Gunlogson 2000; Sudo 2013; Shao 2014). Two recent elaborate accounts on the non-truth-conditional interpretation of PQs are offered by Sudo (2013) and Shao (2014). Sudo (2013) proposes that PQs are exclusively biased, and that to interpret these PQs, both *epistemic bias* (i.e., whether a positive/negative answer is favoured) and *evidential bias* (i.e., whether positive/negative evidence is required) shall be considered. Shao (2014) submits that some PQs are not simply biased *à la* Sudo (2013), but are complex non-canonical questions whose interpretation is constrained by four contextual factors: *epistemic bias* (defined as Sudo's), *antecedent requirement* (i.e., whether a positive/negative antecedent is required), *answer requirement* (i.e., whether an answer is required), and *acceptance status* (i.e., whether the antecedent is likely to be accepted/rejected).

Unfortunately, Sudo's (2013) and Shao's (2014) accounts are both introspective and it remains unclear whether they are empirically valid and which offers a better predication. In this study, these two research questions

are tackled via two short experiments, in which four types of Mandarin positive polar questions (PPQs) marked by either an ending particle (i.e., *ma*, *-ba*, *-a*) or a high boundary tone (H%) were chosen as the testing targets. In view of Sudo's (2013) and Shao's (2014) preliminary analyses, we hypothesized that (i) *evidential bias*, which is not mentioned in Shao (2014), should also play a role in the construal of Mandarin PPQs, and that (ii) Mandarin PPQs or at least some of them are not simple biased PQs *à la* Sudo but are complex non-canonical questions.

## Methods

Two short acceptability judgment experiments were conducted, one aiming to test Sudo's (2013) model, and the other targeting Shao's (2014) proposal.

### Experiment I

Sixty native Mandarin speakers ( $M=25.366$ ,  $SD=1.834$ , Range 20-29) participated in Experiment I. A  $3 \times 3 \times 4$  within-subjects design is adopted with EVIDENTIAL BIAS (3 levels: *neutral*, *positive*, and *negative*), EPISTEMIC BIAS (3 levels: *neutral*, *positive*, and *negative*), and PPQ PATTERNS (4 levels: H%-, *ma*-, *ba*-, and *a*-marked PPQs) set as independent factors. Every four parallel PPQs were tested through a single situation, and for each condition, four tasks were designed. A total of 36 tasks were obtained, 9 of them being fillers. In each task, participants had to read an incomplete monologue/dialogue, preceded by a short context-setting passage, which specified the relevant contextual information. Four corresponding PPQs were later displayed, and the participants' task was to judge whether each of them is suitable (2 levels: *suitable* and *unsuitable*).

### Experiment II

Another group of sixty native Mandarin speakers (Age range 18-26,  $M=21.100$ ,  $SD=2.343$ ) participated in Experiment II. A total of 48 ( $2 \times 2 \times 3 \times 4$ ) judgment tasks (12 of them being fillers) were created, distributed in two groups: Group 1 targeted the correction of POSITIVE ANTECEDENT  $\times$  NEGATIVE EPISTEMIC BIAS but differed with respect to RESPONSE CONDITION (2 levels: *required* and *unrequired*) and ACCEPTANCE STATUS (3 levels: *accepted*, *rejected*, and *neither*); Group 2 targeted the other correlation of NEGATIVE ANTECEDENT  $\times$  POSITIVE EPISTEMIC BIAS, and also differed with respect to the two same factors. The rest of the experimental setting and procedure were similar to that of Experiment I.

### Results

Rating data from both experiments were submitted to several generalized linear mixed models using SPSS 25. The detailed results are spared here due to space limitations, but are retrievable from the following figures:

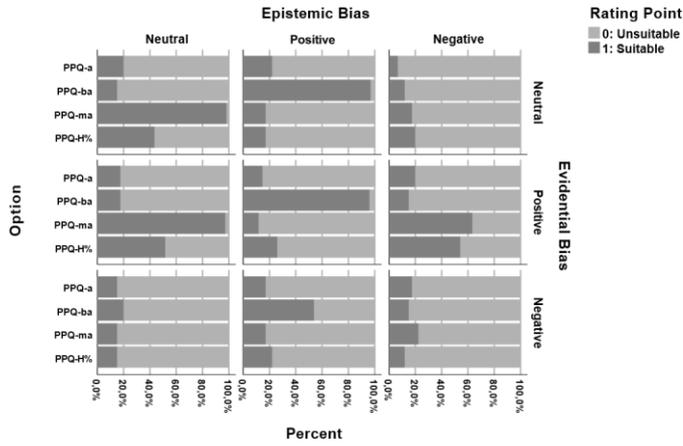


Figure 1. Distribution of data points in different conditions of Experiment I.

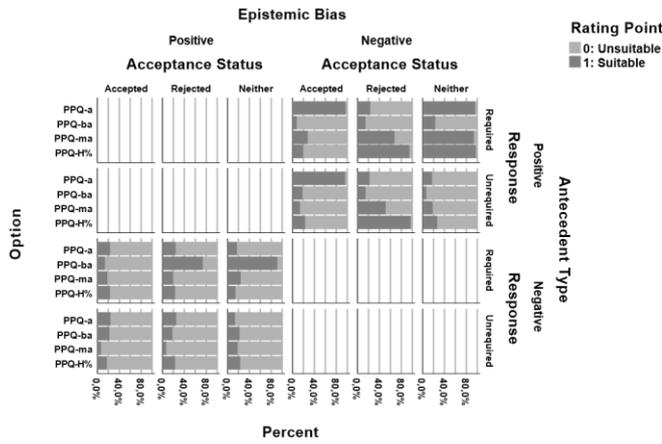


Figure 2. Distribution of data points in different conditions of Experiment II.

### Discussion and conclusions

The results suggest a hybrid picture, as shown in Table 1. In a nutshell, the empirical findings reveal that Mandarin PPQs are not exclusively simple biased PQs – only *ma*- and *ba*-marked PPQs can be characterized in terms of evidential and epistemic biases, and the other two patterns – PPQ-H% and PPQ-*a* – cannot. Interestingly, all of the four patterns can be used to respond to an antecedent. Some require to be answered whereas some others can be

left unaddressed. The acceptance status of antecedent is also relevant for the interpretation of Mandarin PPQs, especially in distinguishing between H%- and *a*-marked PPQs. Five marginal patterns were also reported but are not presented here due to space limitations. A modified feature-based model integrating both Sudo (2013) and Shao (2014) is therefore motivated and is shown to be fine-grained enough to describe Mandarin PPQs.

Table 1. Dominant interpretation patterns of Mandarin PPQs.

Pattern	Evidential	Epistemic	Antecedent	Answer	Acceptance
PPQ-H%	×	+negative	+positive	±	-accepted
PPQ-mal	-negative	-positive&-negative	×	+	-accepted&-rejected
PPQ-ma2	×	+negative	+positive	+	-accepted
PPQ-ba1	-negative	+positive	×	+	-accepted&-rejected
PPQ-ba2	×	+positive	+negative	+	-accepted
PPQ-a	×	+negative	+positive	±	-rejected

(Notation: × indicates that the feature is irrelevant).

## Acknowledgements

This study is supported by four research grants: BES-2016-077488 and FFI2015-67991-P, both awarded by the Spanish Ministry of Economy and Competitiveness, PGC2018-094029-A-100 awarded by the Spanish Ministry of Science, Innovation, and Universities, and 17ZDA027 awarded by the Chinese National Social Science Fund.

## References

- Hamblin, C. L. 1973. Questions in Montague English. *Foundations of Language* 10(1), 41-53.
- Karttunen, L. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1(1), 3-44.
- Büring, D., Gunlogson, C. 2000. Aren't positive and negative polar questions the same. Manuscript, University of California at Santa Cruz.
- Sudo, Y. 2013. Biased polar questions in English and Japanese. In Gutzmann, D., Gärtner, H. (eds.), *Beyond Expressives: Explorations in Use-conditional Meaning*, 275-295. Leiden, Brill.
- Shao, J. 2014. *Study of Interrogative Sentences in Modern Chinese*. Beijing, Commercial Press.
- Ladd, D. R. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In *Papers from the Regional Meeting of Chicago Linguistics Society*, 164-171, Chicago, US.

## Index of names

---

- Alexandersson, J. 133  
 Al-hawi, A. .... 13  
 Alves, J. .... 121  
 Andreou, G. .... 17  
 Antonsson, M. 21, 85,  
 117  
 Barkat-Defradas, M.  
 ..... 101  
 Batista, F. .... 37  
 Bertinotto, P.M. .... 45  
 Besedina, E. .... 25  
 Bleier, R. .... 165  
 Blunier, D. .... 33  
 Botinis, A. .... 29  
 Bourgeois-Gironde,  
 S. .... 33  
 Cabarrão, V. .... 37  
 Camargo, Z. .... 41  
 Cappelli, G. .... 45  
 Cardoso, M. .... 121  
 Cataldo, V. .... 189  
 Chanchaochai, N. .... 65  
 Chernova, D. .... 49  
 Chudar, A. .... 105  
 Collette, C. .... 101  
 Corrêa, L.M.S. 53, 57  
 Costa, A. .... 137  
 Costa, J. .... 61  
 Creemers, A. .... 65  
 Eckerström, ..... 69  
 Eloeva, F. .... 73  
 Embick, D. .... 65  
 Fassetti, F. .... 77  
 Fassetti, I. .... 77  
 Ferrari, L.R. .... 93  
 Ferreira, ..... 61  
 Fontes, M. .... 81  
 Forster, R. .... 57  
 Fraser, K.C. .... 85  
 Frontera, M. ... 89, 161  
 Frouin, C. .... 101  
 Fujiwara, F. .... 97  
 Gayraud, F. .... 101  
 Geurts, B. .... 1  
 Godard, J. .... 185  
 Goritskaya, O. .... 105  
 Gražul-Luft, A. ... 109  
 Hadler Coudry, M.I.  
 ..... 113  
 Harris, R. .... 157  
 Jaeger, T.F. .... 197  
 Jesus, L.M.T. .... 121  
 Johansson, S. .... 117  
 Julião, M. .... 37  
 Kaglik, A. .... 125  
 Karpava, S. .... 13  
 Kisilier, M. .... 73  
 Kokkinakis, D. 21, 85,  
 117, 201  
 Kontostavlaki, A. ... 29  
 Lefebvre, L. .... 101  
 Lenci, A. .... 45  
 Leshchenko, Y. ... 129  
 Li, P. .... 213  
 Liakou, M. .... 17  
 Lindsay, H. .... 133  
 Linz, N. .... 133  
 Loureiro, I.S. .... 101  
 Lourenço-Gomes, M.  
 ..... 209  
 Luegi, P. .... 137  
 Lundholm Fors, K. 21,  
 85, 117  
 Madureira, S. .... 81  
 Maia, M. .... 137  
 Mari, A. .... 33  
 Martin, Ph. .... 141  
 Mata, A.I. .... 37  
 Melebari, A. .... 181  
 Mhanna, E. ... 125, 145  
 Migliaccio, R. .... 125  
 Mograbi, D.C. .... 93  
 Moniz, H. .... 37  
 Morgado, M. .... 121  
 Moutinho, M. .... 113  
 Nakamura, F. .... 97  
 Nasika, F. .... 149  
 Nicolas, D. .... 33  
 Nikolaenkova, O. ... 29,  
 73  
 Nisticò, S. .... 77  
 Ostapenko, T. .... 129  
 Özkayin, G. .... 153  
 Paone, E. .... 89, 161  
 Pereira, ..... 61  
 Pereira, E. .... 165  
 Pham, M.D. .... 157  
 Pimenta, C. .... 169  
 Prudlo, J. .... 133  
 Rákosi, G. .... 173  
 Rilliard, A. .... 41  
 Rodrigues, E. ... 57, 93  
 Ryzhikova, T. .... 177  
 Sabir, M. .... 181  
 Sanches, C. ... 125, 145,  
 185  
 Schettino, L. .... 189  
 Shamina, E. .... 193  
 Solera-Ureña, R. .... 37  
 Stewart, T.C. .... 157  
 Suzuki, D. .... 97  
 Swerts, M. .... 9  
 Tan, M. .... 197  
 Teichmann, M. ... 125,  
 145, 185  
 Themistocleous, Ch.  
 ..... 29, 201  
 Tóth, E. .... 173, 205  
 Trancoso, I. .... 37  
 Tröger, J. .... 133  
 Tsela, V. .... 17  
 Urtegeshev, N. .... 177  
 Valero-Cabré, A. 145,  
 185  
 Wallin, A. .... 69  
 Xie, X. .... 197  
 Yamada, M. .... 209  
 Yuan, Ch. .... 213



ExLing 2019

Ebook ISSN:

Ebook ISBN: