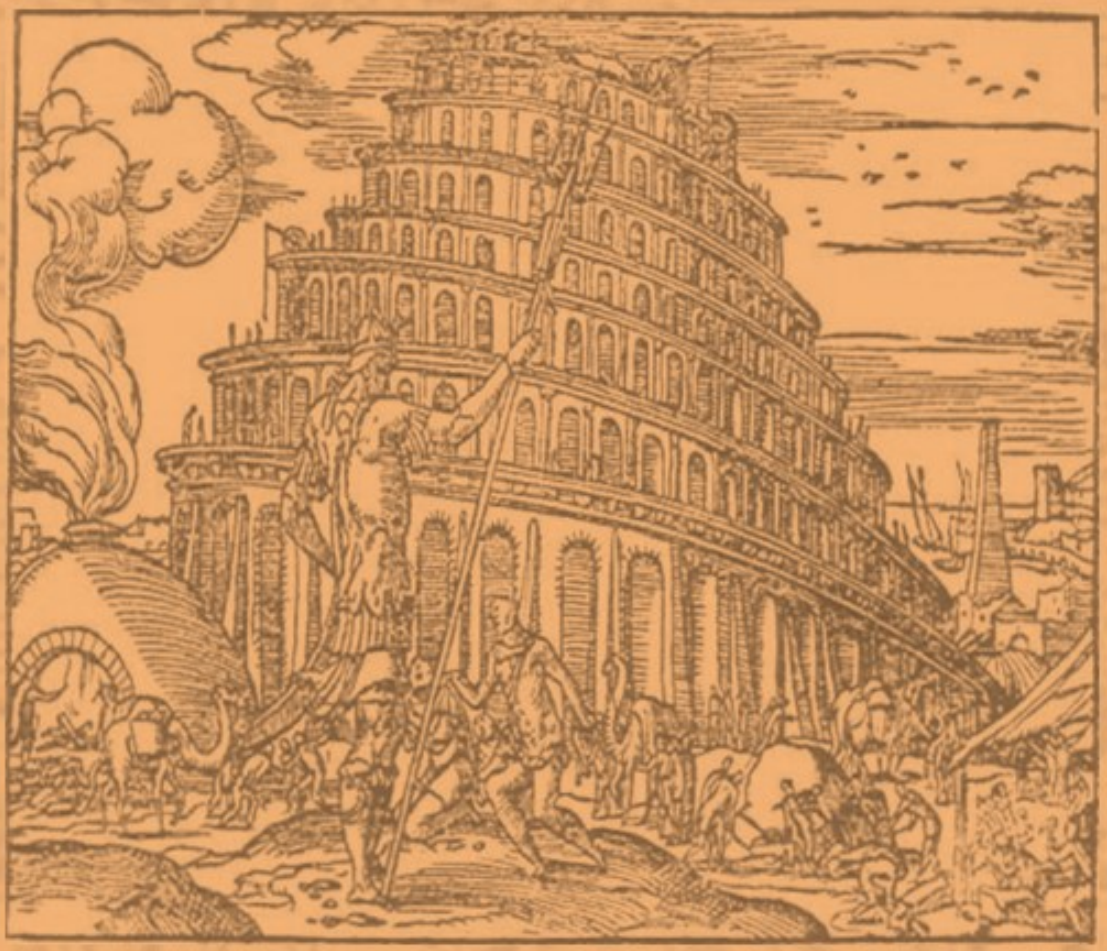
International Society of Experimental Linguistics

ExLing 2020

Proceedings of 11th International Conference of Experimental Linguistics

12-14 October 2020
Athens, Greece

Edited by Antonis Botinis



National and Kapodistrian
University of Athens

International Society of Experimental Linguistics

ExLing 2020

Proceedings of 11th International Conference of

Experimental Linguistics

12-14 October 2020

Athens, Greece

Edited by Antonis Botinis

ExLing 2020
Proceedings of 11th International Conference of Experimental Linguistics

Published by ExLing Society
Electronic edition
ExLing 2020
Athens, Greece
ISSN: 2529-1092
ISBN: 978-618-84585-1-2
DOI: 10.36505/ExLing-2020/11

Foreword

The present volume includes the proceedings of ExLing 2020, the 11th International Conference of Experimental Linguistics. For the first time, we held the conference online, which has unleashed the opportunity for us to host a wider range of events and activities in the future.

Our Society is defined by a collective commitment to understanding language by developing experimental methods in Linguistics. It is an international forum for both established and emerging researchers to participate in and discuss diverse developments in linguistic research and experimental methodologies.

The first Exling Conference – the International Speech Communication Association (ISCA) ExLing Workshop – took place in Athens in 2006. In the years following we have held events across the globe, in cities such as Paris, Saint Petersburg and Lisbon, and, in 2019, established the annual International Conference of Experimental Linguistics.

In that we are hosting it from Athens, ExLing 2020 has a historical connection with our first event 15 years ago. But its virtual format also marks a new chapter, exploring how technology can facilitate the online language interaction of our members without losing the natural interpersonal communication that our Society holds in high regard. To promote this interaction, members of the ExLing Society have free access to the digital conference, thus opening our online presentation to a wider audience.

Alongside the main conference of contributions to various aspects of experimental linguistics and related disciplines, ExLing 2020 has hosted two special sessions encompassing diverse and energetic research domains that focus on current topics in Language Education and Language Pathology.

We would like to thank all ExLing 2020 participants and our keynote speakers, Yosef Grodzinsky, Jonathan Harrington, Kathy Rastle, William Snyder, Wolfram Ziegler, as well as colleagues from the International Advisory Committee, for their contribution to the successful outcome of the Conference.

Antonis Botinis
ExLing Society

Contents

Keynote papers

Research papers

How do writing systems shape reading and reading acquisition?

Kathy Rastle
Department of Psychology, Royal Holloway University of London, UK
https://doi.org/10.36505/ExLing-2020/11/0001/000416

Abstract

Writing is a relatively recent cultural invention, and reading is a skill that requires years of instruction, dedication, and practice. My talk will consider how the nature of a writing system influences reading acquisition and skilled reading. I consider the nature of statistical regularities that characterize English orthography and show across several experiments that knowledge encoded in the skilled reading system mirrors these regularities. This analysis reveals that weaknesses in the relationship between spelling and sound give rise to powerful regularities between spelling and meaning that are critical for text comprehension. I conclude by thinking about how written language differs from spoken language and argue that these differences may be at the heart of human capacity for rapid, skilled reading.

Keywords: writing systems, reading acquisition, language analysis, English

Introduction

By the time that most children leave school, the act of reading seems effortless. Text is all around us and we cannot help but understand what it means. Yet, our experience of reading belies the fact that unlike walking or talking, humans are not born to read. Instead, writing is a relatively recent cultural invention, and reading is a skill whose mastery requires years of instruction, dedication and practice.

When children come to the problem of learning to read, most have already developed substantial knowledge of spoken language. The critical challenge is therefore to map the visual symbols of writing onto this spoken language knowledge, using neural machinery built for other functions. In this talk, I describe how the nature of a writing system shapes this process, and I ask how the evolution of writing may have supported our capacity for skilled reading.

Writing systems

Writing systems always represent spoken language, but they do so in different ways. Visual symbols can represent sounds, syllables, morphemes, or whole words. Much of the research on reading has focused on alphabetic writing systems, in which visual symbols represent sounds. These writing systems can

be categorised further in terms of their orthographic depth, or the consistency with which visual symbols represent sounds.

Research on reading has been dominated by the notion that writing systems that faithfully represent the sounds of language are preferred. If reading is the act of translating visual symbols back to spoken language, then writing systems that offer a direct line back to the phonological forms of words would indeed appear superior. Yet, we accept some deviations between written and spoken language, such as spacing to mark word boundaries, uncritically as a good thing. Do deviations between written and spoken language ever support reading?

English spelling has a bad reputation

Much has been made of the high degree of inconsistency in English spelling: for example, have (cf. gave, wave, save); one (cf. zone, lone, phone); chef (cf. church, chess, chat). Research shows that it takes longer to learn to read aloud in English than it does in other European languages; and learning to spell in English presents an enormous challenge, with many possible spellings for each sound.

Yet, a deeper look suggests that these weaknesses in the English spelling-sound mapping mask another form of regularity. If English were a perfect transcription of spoken language, then words like herded, snored, and kicked might be spelled herdid, snord, and kict. These spellings remove the inconsistency associated with the -ed spelling in the original examples; however, they also remove a powerful morphological cue to the meanings of those words. That is, the letters -ed are almost always associated with the past tense. Other spellings are used in order to prevent incorrect interpretations of words that are not past tense (e.g. evict, not evicked).

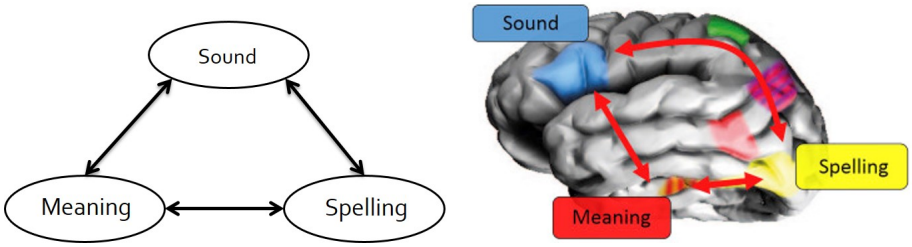
It turns out that this tight relationship between English spelling and meaning is ubiquitous. The many possible spellings for English sound sequences permit some spellings (affixes) to become reserved to communicate particular meanings with a high degree of fidelity. This phenomenon makes it possible to determine rapidly with just a cursory analysis whether a suffixed word is an entity, property, or act. Crucially, this information is only available in the spelling; for example, the spoken forms of evict and kicked both sound like they might be in the past.

This discussion suggests that English spelling might not deserve its bad reputation. The property that makes English spelling hard to learn allows it to communicate meaning with a high degree of precision. This state of affairs turns out to have important consequences for skilled reading.

The reading system is the writing system

Most of the research on reading acquisition in English and in other alphabetic writing systems has focused on how children acquire understanding of the relationship between spellings and sounds. English-speaking children encounter

around 5,000 printed words in the first year of reading instruction, most of which are monosyllabic and have a single morpheme. It is unfeasible to memorize these one-by-one, so instruction is focused on helping children to decode printed words to a phonological code, thus enabling them to make use of their spoken language knowledge to gain access to meaning.



There is ample research to suggest that learning the spelling-sound relationship is a necessary part of reading acquisition. However, several lines of evidence suggest that phonological decoding is not by itself a viable means of driving rapid, skilled reading. The weak relationship between spelling and sound presents a particular challenge in English, as studies of adult readers reveal that the inconsistency that characterises this mapping is mirrored in their linguistic knowledge. Ultimately, skilled reading requires a more efficient, direct mapping between spellings and meanings.

Acquiring this direct mapping presents a challenge because for most short words encountered in the initial stages of reading acquisition, this mapping is arbitrary. However, morphemes provide islands of regularity in this mapping; stems reoccur in words with similar meanings, and affixes alter the meanings of words in highly-predictable ways. Further, because of the trade-off between spelling-sound and spelling-meaning regularity described above, morphology is highly visible in English spelling (much more than in spoken language).

There is substantial evidence to suggest that readers take advantage of this information in building a direct mapping from spelling to meaning. Skilled readers segment morphologically-structured letter strings in the first 200 ms of recognition, in brain regions that underpin the reading pathway linking spelling to meaning. They also show a high degree of sensitivity to the relationship between affixes and aspects of meaning. Adults are more likely to spell the spoken word /dəʊməs/ using -ous if it occurs in an adjective context than in a noun context. Similarly, adults' eye movements are more likely to regress back to the word domous if it occurs in a noun context than in an adjective context. The strength of these effects for different suffixes mirrors the strength of the relationship between those suffixes and grammatical category in English words.

Why do spoken and written language differ?

Writing systems vary in the extent to which they facilitate translation back to a phonological representation. We have seen that the English spelling-sound relationship is relatively opaque. Yet, even those writing systems that offer a faithful transcription of spoken language diverge in important ways: for example, Korean Hangul physically demarcates syllable boundaries; and most writing systems use spacing to denote word boundaries. Why have writing systems evolved or been designed in these ways?

One possibility is that while a transparent spelling-sound relationship facilitates the initial stages of learning to read, phonological decoding by itself is an inefficient means of accessing meaning. It is also important to recognize that translation back to a phonological representation is not the same thing as spoken language: for example, it lacks prosody, gesture, and audio-visual cues to meaning. Rapid computation of meaning may require the orthography to offer information that is not available in the spoken language. This insight becomes very important in thinking about spelling reform or the development of new writing systems.

I've focused on the information available in printed single words, but it is also critically important to recognize that text is very different to spoken language, using richer vocabulary and more complex syntax. Our ability to engage with this weight and complexity of information requires a precision-tuned system for accessing language through vision. Understanding how different forms of writing support the acquisition of that system will provide broader insights into interactions between biology and culture in human cognition.

Further reading

Rastle, K. 2019. EPS mid-career prize lecture 2017: Writing systems, reading, and language. Quarterly Journal of Experimental Psychology, 72, 677-692.

Ulicheva, A., Harvey, H., Aronoff, M., Rastle, K. 2020. Skilled readers' sensitivity to meaningful regularities in English writing. Cognition 195, 1-21.

Berg, K., Aronoff, M. 2017. Self-organization in the spelling of English suffixes: The emergence of culture out of anarchy. Language, 93, 37-64.

Children's syntax: a parametric approach

William Snyder
Department of Linguistics, University of Connecticut, US
https://doi.org/10.36505/ExLing-2020/11/0002/000417

Abstract

Here I present some reasons to take a "parametric" approach to children's acquisition of syntax. I briefly review findings from three case studies which, in my view, offer important insights into what happens when a child's syntax undergoes a change. Each of the case studies is based on longitudinal corpora of spontaneous-speech samples from children acquiring English or Spanish, and each one examines the initial emergence of syntactic structures that are subject to cross-linguistic variation. My principal claims are that changes in a child's syntax are *decisive, additive,* and *interconnected.*

Keywords: child language acquisition, syntax, principles and parameters

Introduction

In the study of child language acquisition, there is a long tradition of focusing on children's errors, a tradition that has unfortunately led to some widespread misconceptions. This tendency may have resulted from a natural sampling bias: whenever a child makes an error of "commission" (i.e., combining morphemes in a way that is disallowed in the target language), alarm bells go off in the mind of any adult observer. Whether it is a child's parent keeping a language diary, or a scientist examining a transcript of a child's speech, the adult's attention will naturally be drawn to the child's errors.

In contrast, when a child assembles morphemes in a way that is actually possible in the target language, there are no mental alarm bells. An adult observer may go for some time without even noticing that a new grammatical structure has begun to occur in the child's speech, if it is something used routinely by adults. Thus, it is quite easy to get the impression that children's acquisition of grammar is a prolonged process of gradual changes, plagued by persistent commission errors along the way.

Yet, when longitudinal transcripts of a child's spontaneous speech are examined quantitatively, a very different picture emerges. In the remainder of this paper I will very briefly present several parts of this picture: the observation that a change in the child's syntax is normally abrupt (or "decisive"); that changes are almost always "additive," in the sense of enlarging the child's existing syntactic repertoire, which is composed entirely of structures (or portions of structures) that are permitted in the target language; and finally that changes can be "interconnected," in the sense that sometimes, certain (seemingly) unrelated structures in a given language are acquired by every child as a package.

Case study I: verb-particle combinations

The first case study (drawn from Snyder 2007, Chapter 4) examines *error patterns* in the speech of a child known as Sarah (Brown 1973), when she was first beginning to use English verb-particle constructions (e.g., rip the lid off). The corpus of recorded conversations between Sarah and her various interlocutors (her friends and family, plus members of Roger Brown's research team at Harvard) was chosen because it was the densest corpus available for a child acquiring English, with a mean of just 7.4 days between recordings that cover an age range from 2 years 3 months (2;03) to 5;01. The English verb-particle construction was chosen as the focus because it is used with high frequency both by adults, and by children once they acquire it; and at the beginning of her corpus, Sarah was not yet using it.

The case study began with a systematic enumeration of the logically possible error-types that a child might make, prior to mastering the adult structure; and the next step was the design of a computer-assisted search that would locate as many as possible of these error-types, if they ever occurred in Sarah's corpus. The list of possible error-types included errors expected from a child who was using analogical reasoning (e.g., * She picked up it, on analogy with sentences like She picked [the book] up, She picked [it] up, and She picked up [the book]). It also included errors expected from a child who was considering possible grammars that are close to (but distinct from) the actual grammar of English, such as the grammars of languages that are typologically similar (and historically related) to English. Errors of this type would include, for example, She has the book up-lifted, on the model of the word-for-word counterparts that are possible in Dutch and German.

The main finding was that Sarah made an abrupt, *"decisive"* change (at the age of 2;06) from producing extremely few (if any) genuine verb-particle combinations, to suddenly producing verb-particle combinations comparable to those of adult English. Moreover, right from the outset, Sarah's verb-particle combinations were overwhelmingly correct, from the perspective of the adult's grammar for English. Almost all the errors that did occur were errors of *omission, not commission;* and these errors did not begin to occur until the same point when the correct forms had also begun to occur.

Case study II: prepositional questions

The second case study (from Sugisaki & Snyder 2006) concerns acquisition of prepositional questions ('P-questions') in English and Spanish. In English, the P is normally "stranded" (e.g., *What* was Peter talking *about*?), while in Spanish the P is "pied-piped" along with the wh-expression (e.g. *De qué* hablaba Pedro?, lit. "[About what] was-talking Peter?"). What if a child wants to ask a P-question and does not yet know how? Is there a "default" way? (The English way? Spanish? Something else?) To find out, we examined ten longitudinal corpora for English, and four for Spanish. Our computer-assisted searches located

every child utterance that contained both a P and a word or phrase capable of serving as a wh-expression.

What we discovered was that children *never used a "default" option*. When children acquiring English first began producing P-questions, they used P-stranding, just like adults. When children acquiring Spanish first began producing P-questions, they used pied-piping, just like adults.

What was even more striking was that four of the children had a quite sizable gap (range: 2.0-9.0 months; mean 5.2) between the point when they were producing both direct-object (DO) questions and the declarative counterparts to P-questions; and the point when they began producing P-questions. In other words, these children went for up to *9 months* without asking P-questions, even though they produced numerous DO-questions (range: 11-48; mean: 29.8) during the gap.

Moreover, once the children began producing P-questions, the P-questions were immediately used almost as frequently as DO-questions. Hence, it appears that a child who does not yet know how to produce a P-question correctly in her target language actually refrains from even attempting to ask such questions in spontaneous speech. This suggests the child temporarily has a grammar providing the correct structure for DO-questions but providing no structure at all for P-questions; and that a subsequent syntactic change *"adds"* P-questions to the grammar.

Case study III: compounds and complex predicates

The final case study (from Snyder 2007, Chapter 5) shows that acquisition can be *"interconnected."* In (Snyder 1995) I proposed the existence of the "Compounding Parameter," a point of cross-linguistic variation that affects both compound words and complex predicates. The key idea is that certain languages (the ones with the positive setting of the parameter, such as English) freely allow the creation of novel, bare-stem, compound words (e.g., animal cup, for a cup with pictures of animals on it; or animal cup catalog, for a catalog offering different types of animal cup), whenever the context is sufficient to make the intended meaning clear. Languages with this setting include all the Germanic languages, plus Hungarian, Mandarin, Japanese, and Thai. Languages with the negative setting include all the major languages of the Romance and Slavic families. In such languages the ideas that are normally expressed with a compound word in English generally have to be expressed in phrasal syntax, using something like a preposition or an oblique case to create a modifier.

Moreover, several types of complex predicate seem to be found only in languages with the positive setting of the Compounding Parameter. One example is the separable verb-particle construction examined in Case Study I; this is a complex predicate in the sense that the verb and the particle jointly determine the event type (Aktionsart) of the verb phrase. For example, a simple activity like carrying a box – John carried the box (?? in thirty seconds) – can be

converted to an accomplishment through the addition of a particle – John carried the box out in thirty seconds – even with the particle is separated from the verb by a full-DP direct object.

Proposing an explanation in terms of an abstract grammatical parameter, within a Principles-and-Parameters framework, leads to a very strong prediction for child language acquisition: in children acquiring a language like English, with the positive setting, the point in time when any given child acquires the verb-particle construction should be tightly correlated with the point in time when the same child acquires bare-stem compounding. In (Snyder 2007) I reported the results from a case study that tested this prediction directly, using 19 high-quality longitudinal corpora from the CHILDES database (MacWhinney 2000). Age of acquisition was measured as age of the first clear use, provided that use was soon followed by additional uses with different lexical items. The children varied considerably in the ages at which they acquired verb-particle combinations, and they also varied considerably in the ages of acquisition for bare-stem compounding. Yet, for any single child, the ages of onset for verb-particle constructions and novel N-N compounding were nearly identical; this yielded a highly robust statistical correlation ($r=0.937$, $t(17)=11.1$, $p<.0001$). Hence, the findings provided striking support for a parametric approach.

Discussion and conclusions

Taken together, the observations that syntactic changes are decisive, additive, and interconnected support a view of language acquisition in which the child is collecting evidence about the target language, and determining the values of grammatical parameters, with extraordinary precision.

References

Brown, R. 1973. A first language: The early stages. Cambridge, MA: Harvard University Press.

MacWhinney, B. 2000. The CHILDES Project (3rd edition). Volume I: tools for analyzing talk. Mahwah, NJ: Lawrence Erlbaum.

Snyder, W. 2007. Child language: the parametric approach. Oxford University Press.

Sugisaki, K., Snyder. W. 2006. The parameter of preposition stranding: a view from child English. Language acquisition 13:4, 349-361.

A neurophonetic perspective on articulation planning[1]

Wolfram Ziegler
EKN, Institute of Phonetics, Ludwig-Maximilians-University, Munich, Germany
https://doi.org/10.36505/ExLing-2020/11/0003/000418

Abstract

This paper gives an overview of a model that predicts articulation ease for German phonological words on the basis of error data from patients with apraxia of speech (AOS). AOS is introduced as a clinical model of higher order motor processes for articulation. Word production accuracy in AOS is considered as a window into the structure of articulation plans as acquired through speech motor learning in childhood. The NLG model of apraxia of speech is explained. Applications in speech development and adult speech are outlined.

Keywords: articulation apraxia of speech

Neurophonetics

Neurophonetics is a research field that uses the knowledge and methodological tools developed in phonetics to study the neurological aspects of speaking. It covers investigations of the impact of brain dysfunctions on speech production in patients with neurologic conditions, but also of the neural substrates of speech in typical speakers.

Two different lines of neurophonetic research can be distinguished. First and foremost, phonetic thinking and methodologies are applied in the service of clinical neurology. An increasing number of studies is devoted to describing the speech patterns of patients with neurologic disorders, such as stroke, Parkinson's disease, or cerebellar ataxia. This research mainly aims at understanding the impact of neural dysfunction of specific aetiologies or localizations on the patients' speech characteristics. Applications of this research seek to establish physiologic, acoustic, or auditory-perceptual parameters that are sensitive and specific to neurologic speech impairments, with the ultimate goal of developing phonetically based clinical assessment tools (cf. Duffy 2019).

A second, less frequented line of neurophonetic research goes the opposite way, aiming to uncover general principles of speech production by investigating their breakdown due to dysfunctions of relevant brain networks. The study of disordered cognitive functions to learn more about the neural organization of "normal" cognition has long been a productive principle in cognitive neuropsychology, especially in neurolinguistics, though the "transparency" assumption underlying this principle has always been disputed. This assumption

claims that the impairment resulting from brain lesions is transparent for the structure of the underlying, unimpaired cognitive functions (Caramazza 1988). As a major precondition, this approach must be based on the solid foundations of a clinical model whose relationship to the cognitive function in question is sufficiently well established.

In this paper I will present a well-established clinical model of higher-order speech motor functions, *apraxia of speech* (*AOS*), and describe how investigations of this model may create evidence about the organization of typical speech. The disorder is considered to result from a dysfunction of the ability to plan the articulator movements required to produce the syllables, words, and sentences of the patient's native language.

A clinical model of articulation planning

Apraxia of speech occurs predominantly in patients who have suffered a left hemisphere stroke. Initially the patients are often completely unable to articulate, but within several hours or days their speech gradually recovers. Their articulation is usually slow, dysfluent, and effortful, and they produce errors such as substitutions, omissions and distortions of speech sounds, schwa intrusions in consonant clusters or cluster reductions, slowed or distorted transitions between speech sounds, and impaired coarticulation. They show visibly laborious groping movements of the articulators, with repeated attempts, false starts, and restarts in the initiation of utterances. The syndrome may manifest itself over a wide range of severity levels, from almost complete mutism to only mildly dysfluent speech with occasional articulation errors, and with different recovery dynamics across patients. For descriptions of the clinical pattern of AOS see e.g. Duffy (2019) or Ziegler (2008).

An important issue in the theoretical classification of this disorder is to differentiate it from other types of neurogenic sound production impairment. On the one hand, the symptoms of AOS are not explainable by "elementary" motor pathologies of the vocal tract muscles, such as paresis, ataxia, hypo- or hyperkinesia etc., which are summarized under the clinical term *dysarthria*. Unlike the dysarthrias, which are typically caused by bilateral brain lesions, AOS is a syndrome of the language-dominant hemisphere. The label "apraxia" historically relates to exactly this circumstance and characterizes the disorder as an impairment of "higher" motor functions. On the other hand, the syndrome is different from aphasic-phonological impairment, which is characterized by phoneme errors in essentially fluent speech, without any apparent signs of articulomotor involvement. Though these differential diagnostic considerations are far from trivial and have often been disputed, AOS is widely accepted as an autonomous clinical unit, and an enormous amount of research has been devoted to its clinical pattern and its neuro-anatomic substrate. For discussions see Ziegler, Aichert, and Staiger (2012).

Research interest in this condition has existed for almost 160 years. Broca's seminal case study of a man who lost "la faculté du langage articulé" after a lesion to the posterior part of the left inferior frontal gyrus (Broca, 1861) is generally acknowledged as the birth of systematic clinical brain-behaviour research. Broca pointed out, in the terminology of his time, that the patient was unable to articulate, although he had no obvious motor restrictions of the tongue and lips and no generalized language or cognitive impairment. He allocated the ability to articulate to the left frontal cortical region that bears his name until today, and characterized the speech impairment as a "loss of the memory of the procedures required for the articulation of words" (Broca 1861; p. 333).

Since Broca's time, the neuro-anatomic basis and the functional characterization of this condition have been discussed extensively, - probably more than any other neurologic speech or language dysfunction. Over decades, evidence has accumulated that lesions to left posterior inferior frontal cortex including the opercular part of Broca's area and the adjacent pre-motor and motor cortex, as well as subjacent anterior insular cortex, are responsible for the development of AOS. Moreover, in more recent years this region has been identified in numerous imaging studies of non-brain-damaged speakers as a higher-order speech motor centre. The dysfunction resulting from lesions to this cortical area has been characterized in varying terms as "apraxia of the language muscles", "phonetic disintegration of speech", a "programming deficit", a breakdown of the "functional coalitions" of articulation, or, more recently, a disorder of "phonetic planning" or "speech motor planning" (for a historical review see Ziegler et al. 2012).

Implied in this thinking is that the pathomechanism of AOS disrupts the language-specific motor patterns for the production of syllables, words and sentences as acquired during speech development. We have recently proposed a model that delineates how speech motor learning in childhood is mediated by subcortical structures and leads to an accumulation of a "knowledge base for articulation" in the left frontal cortex of the adult brain (Ziegler & Ackermann 2017). An obvious assumption, based on a wealth of knowledge about the learning-dependent plasticity of the human brain, is that speech motor patterns that are more strongly integrated through speech learning are represented more redundantly within the functional network of this brain region. By implication, such patterns are less vulnerable to a loss of cortical tissue in this area. Conversely, articulatory patterns that are less typical of the speaker's native language, or less cohesive, have less redundant neural representations and are therefore more vulnerable to cortical damage. For a more detailed explanation of this argument see Ziegler, Lehner, Pfab & Aichert (2020).

As a conclusion, the speech patterns of patients with AOS provide a window into the make-up of the acquired implicit knowledge about how the words and sentences of their language are articulated. The argument is similar to that

brought up in studies of speech errors in healthy speakers (e.g., Pouplier & Hardcastle 2005), but with the difference that a much larger corpus of data can be acquired from patients with AOS under relatively natural speaking conditions, and that the errors made by patients with AOS can be allocated to a rather circumscribed functional and neuro-anatomic source, i.e., a dysfunction of the acquired articulation planning processes located in the left posterior inferior frontal lobe.

Accuracy of apraxic word production as a yardstick of "articulatory ease"

For the reasons mentioned above, the errors made by patients with AOS reveal what is easy and what is difficult to articulate for adult native speakers of a given language. "Ease of articulation" is a disputed concept, at least for adult speech, because the highly overlearned nature of the articulation patterns of our native language entails that everything is equally easy for us to say (Ladefoged 1990). In the sense used here, a word is easy to pronounce if patients with apraxia of speech have relatively few problems producing it. More specifically, patients with only mild impairment may mostly produce it correctly, and only those with severe AOS make errors on it. Conversely, difficult phonetic patterns are those that provoke errors even in patients with mild AOS. Following the transparency assumption explained above, the susceptibility of words to apraxic failure mirrors the redundancy of the representation of its motor components in left inferior frontal cortex, which, in turn, is considered the neural substrate of the degree to which language-specific articulatory patterns were stabilized through speech motor learning during childhood.

In many of the earlier studies of AOS, error rates were related to phonemes, mostly with the finding that consonants are less error prone than vowels, plosives and nasals less than fricatives or affricates, voiceless obstruents less than voiced obstruents, etc. Other findings were related to syllable structure, e.g., that coda consonant errors are less frequent than onset errors, or simplex syllables are less vulnerable than complex syllables. Finally, at the supra-syllabic level a rather common finding was that the likelihood of an error increases with the number of syllables in a word. More recently, we could also demonstrate an effect of lexical stress, showing that in German AOS patients, trochaic words were "easier" to pronounce than iambic words (Aichert, Späth & Ziegler 2016). An overview of these findings including the corresponding references is listed in Table 1 of Ziegler et al. (2020).

The NLG-model

Considering that the hierarchies of articulatory planning requirements sketched above extend across all levels of the phonological architecture of words, an approach that integrates these levels is necessary to account for the interactions

between them. The Nonlinear Gestural (NLG) model of AOS, which is based on ideas from Articulatory Phonology (e.g., Goldstein, Pouplier, Chen, Saltzman, & Byrd 2007) considers articulatory gestures as the basic elements of articulation planning and postulates that these units are successively integrated to form syllable structure components, syllables, metrical feet, and phonological words. The model has been described in more detail elsewhere (Ziegler & Aichert 2015; Ziegler, Aichert, & Staiger 2017; Ziegler et al. 2020). In brief, it postulates that in a speaker with AOS of a given degree of severity, the likelihood that an articulatory gesture is planned accurately assumes a certain value $p \in]0,1[$. The more severe the disorder, the more likely is that the production of a gesture fails, hence the smaller is p. By combinatorial laws, the probability that a combination of two gestures is correct drops to p^2, but only under the unlikely assumption that they are independent. More generally, the likelihood that two combined gestures are accurate is expressed as $p^2 * c$, with an unknown coefficient c indicating whether the probability of accurate production increases ($c>1$) or decreases ($c<1$). An increase would point at some inherent bonding of the two gestures, either by biomechanical conditions or through a learned gestural integration, which makes them less vulnerable because they do not count as two independent occasions to fail. Conversely, a decrease would suggest that putting the two gestures together requires additional effort, which makes them more vulnerable than predicted from purely combinatorial considerations. Hence, the coefficients in the model represent the "glue" between the gestural components of syllables and words. Depending on the structural relationship of a gesture with other gestures in a word, different coefficients are inserted to represent the specific degree of integration of the respective gesture within the word's gestural architecture.

Figure 1 illustrates, as an example, the gestural score of the three-syllabic word *Trompete* (engl. *trumpet*). The figure shows a variety of inter-gestural relationships implemented in the NLG-model, such as synchrony of a vocal tract gesture with a glottal (e.g., /t/) or velar gesture (e.g., /m/), consonantal gestures in pre- or post-vocalic positions, gestures clustered within a pre-or post-vocalic position, gestures in the tail of a metrical foot (/tə/) attached to those in the head (/pe:/), and those of an extrametrical syllable to the gestures within a metrical foot. Computationally, each word is modelled by a nonlinear, multiplicative term representing each gesture's base probability of correct production and the coefficients coding for the gesture's bonding strengths within the word's gestural score, as displayed in figure 1.

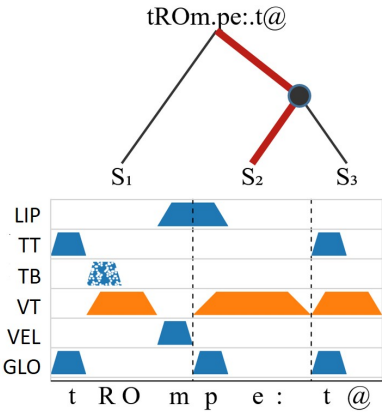


Figure 1: Gestural score of the 3-syllabic German word Trompete (with SAMPA transcript). Each articulatory gesture is represented by a trapezoid. Consonant gestures are arranged on different layers depending on the involved constriction organ (LIP: lips, TT: tongue tip, TB: tongue back, VEL: velar aperture, GLO: glottal aperture). Closed constrictions are represented by filled trapezoids, the "critical" constriction type of the /ʁ/ gesture by a patchy trapezoid. Vowel gestures (in yellow) are allocated to a separate layer representing the gross vocal tract deformations of vowel articulation, including jaw, tongue body and lip rounding movements. Dashed vertical lines indicate syllable boundaries, the tree structure characterizes the word's stress pattern (thick lines indicating prominence).

Model coefficient estimates were calculated and validated in nonlinear regression analyses using corpora of 96 and 136 words of different lengths and complexities, with 120 and 66 renditions per word, respectively, by patients with AOS (Ziegler & Aichert 2015; Ziegler et al. 2020).

Model shape

The sizes of the model coefficients indicate where particularly strong or particularly weak cohesions exist within the gestural patterns of words. The results largely corroborate earlier findings about the factors influencing accuracy of apraxic speech, but unlike earlier investigations of isolated factors, the data present an integrative view that takes account of the interactions between the hierarchical levels of the model.

Here is a selection of findings: onset gestures are more demanding than coda gestures, glottal and velar aperture gestures are highly cohesive with their synchronous tongue- or lip gestures, cluster gestures are independent, gestures within the tail-position of a metrical foot are inexpensive, whereas left-branching expansions of metrical feet are expensive (Ziegler et al. 2020). A transfer of these results to the motor organization of typical, neurologically

healthy speech is only possible on a relative scale. For example, the finding that the coefficient representing the cohesive ties of two cluster gestures is close to 1 suggests an independent recruitment of both gestures, i.e., no integration, in patients with AOS. This does not preclude that gestures in consonant clusters are organized as cohesive units in healthy speakers (Hoole & Pouplier 2015). However, the ties between a tongue back closure and glottal aperture gesture in the production of /k/, with a coefficient considerably greater than 1, are stronger than those between the same tongue back gesture and a tongue tip closure in the cluster /kn/.

Applications

Beyond the clinical applications of the NLG model (Ziegler et al. 2020), this research offers options for investigations of typical speech in adults or in children. A web application for calculating an NLG-score for each word or nonword that meets German phonotactic constraints can be accessed at https://neurophonetik.de/sprechapraxie-gestenmodell. This tool allows for the computation of NLG scores for large lexical databases, (e.g., SUBTLEX-DE, a German lexicon containing word frequency data, https://neurophonetik.de/subtlex-np), with options to investigate the relationships between articulatory ease on the one hand and lexical frequency, lexical familiarity, neighbourhood density, or age of acquisition on the other (e.g., Lehner & Ziegler, accepted). Furthermore, corpora of child word forms can be examined for their relationships with their corresponding adult forms, to investigate hypotheses concerning the role of articulation in infant word learning. Examples of such data will be included in the oral presentation.

Acknowledgement

This research was supported by grants from the German Research Council (DFG, Zi 469 / 6-1/8-1/14-1,2). I am grateful to Ingrid Aichert and Anja Staiger for their long-standing valuable contributions to this work and to the many patients who have agreed to participate in our research.

Notes

1. An expanded version of this paper was submitted to Laboratory Phonology An expanded version of this paper was submitted to Laboratory Phonology.
2. See www.neurophonetik.de/gesten-koeffizientenrechner.

References

Aichert, I., Späth, M., Ziegler, W. 2016. The role of metrical information in apraxia of speech. Perceptual and acoustic analyses of word stress. Neuropsychologia 82, 171-178.

Broca, P. 1861. Remarques sur le siège de la faculté du langage articulé; suivies d'une observation d'aphémie (perte de la parole). Bulletins et memoires de la Société Anatomique de Paris, XXXVI 330-357.

Caramazza, A. 1988. Some aspects of language processing revealed through the analysis of acquired aphasia: The Lexical System. Ann Rev Neurosci 11, 395-421.

Duffy, J. 2019. Motor Speech Disorders: Substrates, Differential Diagnosis, and Management: Elsevier Health Sciences.

Goldstein, L., Pouplier, M., Chen, L., Saltzman, E., Byrd, D. 2007. Dynamic action units slip in speech production errors. Cognition 103, 386-412.

Hoole, P., Pouplier, M. 2015. Interarticulatory Coordination. Speech Sounds. In M. A. Redford (ed.), The Handbook of Speech Production, 133-157. Chichester: Wiley.

Ladefoged, P. 1990. Some reflections on the IPA. J Phonet 18, 335-346.

Lehner, K., Ziegler, W. accepted. The impact of lexical and articulatory factors in the automatic selection of test materials for a web-based assessment of intelligibility in dysarthria. J Speech Lang Hear Res

Pouplier, M., Hardcastle, W. 2005. A re-evaluation of the nature of speech errors in normal and disordered speakers. Phonetica 622, 227-243.

Ziegler, W. 2008. Apraxia of speech. In G. Goldenberg M. Miller (eds.), Handbook of Clinical Neurology , 269-285. London: Elsevier.

Ziegler, W., Ackermann, H. 2017. Subcortical Contributions to Motor Speech: Phylogenetic, Developmental, Clinical. Trends Neurosci 40, 458-468.

Ziegler, W., Aichert, I. 2015. How much is a word? Predicting ease of articulation planning from apraxic speech error patterns. Cortex, 69, 24-39.

Ziegler, W., Aichert, I., Staiger, A. 2012. Apraxia of speech: Concepts and controversies. Jf Speech Lang Hear Res 55(5), S1485-S1501.

Ziegler, W., Aichert, I., Staiger, A. 2017. When words don't come easily: A latent trait analysis of impaired speech motor planning in patients with apraxia of speech. J Phonet 64, 145-155.

Ziegler, W., Lehner, K., Pfab, J., Aichert, I. 2020. The nonlinear gestural model of speech apraxia: clinical implications and applications. Aphasiology, 1-23. doi:10.1080/02687038.2020.1727839

Masked priming in picture naming and lexical selection

Manal Alharbi
School of English, Communication & Philosophy, Cardiff University, US
https://doi.org/10.36505/ExLing-2020/11/0004/000419

Abstract

Our study focused on the co-activation of lexical nodes in bilinguals' speech production. We investigated whether the co-activated lexical forms compete for selection or not and whether language proficiency level would modulate the co-activation level in bilinguals. We tested the performance of Arabic-English bilinguals using the masked priming paradigm in a picture naming task. We found that the co-activated lexical forms do not compete for selection and that proficiency level does not affect the process of lexical selection.

Keywords: lexical selection, lexical access, speech production, different script

Introduction

Most theories of bilingual lexical access in speech production (Costa et al. 1999, Kroll and Stewart 1994, Poulisse and Bongaerts 1994) assume that lexical nodes in the target and non-target language are activated. Hence the question: if various nodes are activated how does the bilingual speaker select the target word? Two conflicting views emerged, namely the language non-specific view and the language specific view (for review, see Costa 2006). The language non-specific view assumes that lexical nodes in both languages are active and compete for selection, and there is an inhibitory mechanism (Inhibitory Control IC) that suppresses the activation of the lexical nodes in the non-target language (Green 1998, Hermans et al. 1998). On the other hand, the language specific view assumes that the lexical nodes in the target and non-target language are active, but only those in the target language are considered for lexical selection (Costa and Caramazza 1999, Costa et al. 1999).

The masked priming paradigm can be a suitable method in testing these theories. Inspired by Costa and Caramazza (1999), we used the identity condition to gain insights into the manner of language selection[1]. If pictures preceded by related masked primes in L1 were named faster than those preceded by unrelated masked primes, this would give evidence to the language specific view. In contrast, if longer naming latencies were found, this would give evidence to the language non-specific view as it would suggest competition between lexical nodes at the lexical level.

Method

Most studies on lexical access in bilingual speech production have tested bilinguals who speak languages that use the same script (the Roman alphabet). To the best of the author's knowledge, only a few studies have investigated the selectivity issue among bilinguals whose two languages have different scripts (Hoshino 2006, Jiang 2012). Therefore, to address the disagreement in the literature, we tested different script bilinguals (Arabic-English bilinguals) and we adopted the masked priming paradigm in picture naming as a new tool of investigation.

Participants

Sixty-Five Arabic-English bilinguals volunteered to participate and they were divided into two groups according to their level of proficiency.

Materials

Thirty white line-drawing picture of objects with non-cognate names were chosen. Two types of prime words were selected for each picture: (1) a prime word in L1 consisting of the translation word for the name of the picture, and 2) a prime word in L1 unrelated to the picture name. For example, a non-cognate picture of a bicycle was primed with the name of that picture in L1, i.e. the translation equivalent (دراجه- bicycle) in the related condition and with an unrelated prime (عنب – grapes) in L1 in the unrelated condition. Each participant was presented with 30 prime-target pairs (15 pictures with related primes and 15 with unrelated primes). The stimuli were counterbalanced across the two different prime conditions.

Design and procedures

The experiment was of a mixed design (i.e. a 2 x 2 design) with proficiency as the between subject factor and the prime type as the within the subject factors. Participants were tested individually and placed in front of high performing multimedia personal laptop, which had a 15" FHD screen and a resolution of 1920 x 1080 and the picture size was 300 x 300 cm. The responses of the subjects were recorded using an external electric microphone connected to the response device Chronos, which measures Response Times (RT). The experiment was designed using E-Prime 3.0 (i.e. a psychology software). Participants were informed that they were required to look at the centre of the screen and name the pictures as rapidly and accurately as possible, and to say 'pass' if they did not know the name of the picture. They were not informed of the presence of the primes. Each trial consisted of the following sequence of events: (1) the tests initially commenced with a fixation point (+), which appeared in the middle of the screen for 500 ms; (2) the appearance of a visual mask of (#####) symbols replaced the fixation point, which remained for 500 ms; (3) the appearance of a prime word (related or unrelated) for 50 ms on the screen; (4) a visual mask of (#####) symbols, which again appeared in the

centre of the screen for 14 ms; and (5) a target picture, which appeared and remained on the screen until the participants made a response.

Results

Reaction times (RTs) were calculated only for correct responses and those longer than the mean plus two standard deviations for a given subject in all conditions were not included in the analysis. A mixed ANOVA analysis was performed on the mean response latencies per subject with: a) proficiency level as a between-subject factor, and b) the relatedness nature of the prime word (related and unrelated prime words) as within-subject factors. A statistically significant main effect was found for the type of prime in both the analysis by subject F1 (1, 26) = 23.851, $p < 0.05$ and the analysis by item F2 (1, 29) = 5.145, $p < 0.05$. Pictures were named faster when they were preceded by masked related prime relative to unrelated prime. An insignificant interaction was found between the type of prime and proficiency level in both the analysis by subject F1 (1,26) = 0.505, $p > 0.05$ and the analysis by item F2 (1,29) = 3.7, $p > 0.05$ (figure 1). There was no significant main effect of proficiency level on the overall scores of reaction times in the analysis by subject F1(1,26) =.026, p >.05 and in the analysis by item F2(1,29) =.064, p>.05 (figure 2).

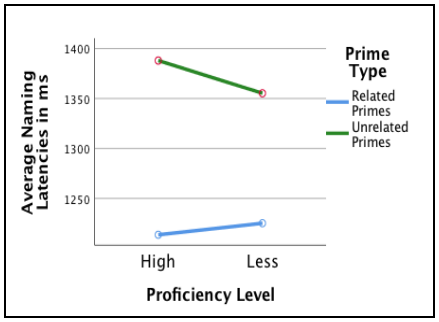


Figure 1. The interaction between prime type and proficiency level.

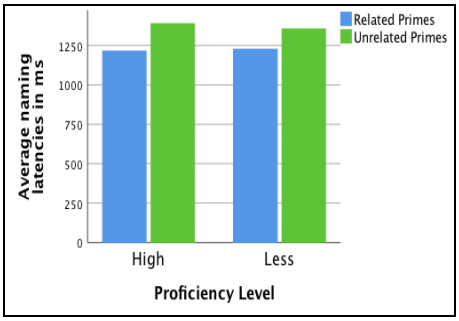


Figure 2: Reaction times in the four conditions according to proficiency level.

General discussion

The purpose of this experiment was to investigate the lexical selection mechanism in Arabic-English adult bilinguals to determine whether it is language-specific or language non-specific and to investigate whether language proficiency level would affect the manner of language selection.

With regard to the first question, it was hypothesized that if the selection mechanism was language-specific, a facilitation effect in this critical condition would be noted. In contrast, if the selection mechanism was language non-

specific (selection by competition), an interference effect in the identity condition would be found. The results showed that naming non-cognate pictures in L2 produced a facilitation effect when preceded by masked related primes in the L1. Pictures preceded by related primes were named 152ms faster than those preceded by unrelated primes. Moreover, the performance of both highly and less proficient were similar in all conditions. The results thus demonstrated that lexical nodes in both target and non-target languages were active at the lexical level and did not compete for selection as no interference effect was observed. Also, proficiency level did not affect the manner of language selection.

Notes

1. The identity condition refers to the condition during which the distractor word is the translation equivalent of the picture to be named in a picture interference task (Costa et al., 2000).

References

Costa, A. 2006. Speech Production in Bilinguals. In: Bhatia, T., Ritchie, W. (eds.) The Handbook of Bilingualism.

Costa, A., Caramazza, A. 1999. Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. Bilingualism: Language and Cognition 2, 231-244.

Costa, A., Miozzo, M., Caramazza, A. 1999. Lexical Selection in Bilinguals: Do Words in the Bilingual's Two Lexicons Compete for Selection? Journal of Memory and Language 41, 365-397.

Green, D. 1998. Mental control of the bilingual lexico-semantic system. Bilingualism: Language and Cognition 1, 67-81.

Hermans, D., Bongaerts, T., De Bot, K., Schreuder, R. 1998. Producing words in a foreign language: Can speakers prevent interference from their first language? Bilingualism: Language and Cognition 1, 213-229.

Kroll, J.F., Stewart, E. 1994. Category Interference in Translation and Picture Naming - Evidence for Asymmetric Connections between Bilingual Memory Representations. Journal of Memory and Language 33, 149-174.

Poulisse, N., Bongaerts, T. 1994. First Language Use in Second Language Production. Applied Linguistics 15, 36-57.

Hoshino, N. 2006. A psycholinguistic study of native language constraints on speaking words in a second language. Pennsylvania State University.

Moon, J., Jiang, N. 2012. Non-selective lexical access in different-script bilinguals. Bilingualism-Language and Cognition 15(1), pp. 173-180.

Syllable rate vs. segment rate in perceived speech rate

Yahya Aldholmi
Department of Linguistics, King Saud University, Saudi Arabia
https://doi.org/10.36505/ExLing-2020/11/0005/000420

Abstract

This study examines the impact of differences in segment rate, which do not correspond to differences in syllable rate, on perceived speech rate in an unfamiliar language. Japanese speakers who were unfamiliar with Arabic judged the speech rate of Arabic stimuli with two levels of syllable complexity on a 1-7 scale. The participants perceived the complex syllable stimuli as being faster than the simple syllable stimuli. This is interpreted as evidence that differences in segment rate *do* influence speech rate perception. Hence, the significance of syllable rate, rather than of segment rate, for perceiving speech rate in some previous studies was likely due to extraneous factors associated with the stimuli and participants in those studies.

Keywords: syllable rate, segment rate, syllable complexity, speech rate, perception

Introduction

Previous studies have implemented various metrics to quantify speech rate, with syllable rate being the most commonly used (e.g., Aldholmi, Park, 2019), followed by segment rate (e.g., Seifart et al., 2018), among other measures. Two utterances with nearly the same syllable rate may have two different segment rates, and vice versa, due to differences in the respective languages' syllable structures. For instance, in the two English and Japanese six-segment words /stɪkt/ "strict" and /kisoku/ "rules", respectively, English incorporates all six segments into one syllable and hence exhibits a high segment rate *per syllable* (i.e., syllable complexity), while Japanese distributes the six segments over three syllables and hence exhibits a lower segment rate. Assuming speech duration is roughly the same (500s sec) for the two words, the speech rate would be 12 seg/sec (or 83.33 sec/seg) for both, but 2 syl/sec (or 500 sec/syl) for the English word and 6 syl/sec (or 166.66 sec/syl) for the Japanese word.

Considering such discrepancies, a few recent studies have attempted to *directly* investigate the impact of syllable complexity on perceived speech rate with varying results. For instance, syllable rate substantially affects speech rate perception for English natives (e.g., Plug & Smith, 2018), while segment rate (i.e., mora count) impacts speech rate perception for Finnish natives (e.g., O'Dell & Nieminen, 2019). More recently, segment rate was found to be related to stimulus duration in English. That is, "among stretches of speech

with various durations… stretches with higher segment rates… are judged as faster than stretches with lower segment rates" (Plug et al, 2020, p. 9).

Language familiarity may have played a role in these findings, because both listener-related factors such as phoneme restoration in language listening and language-specific characteristics such as syllable structure and mora rhythm could influence listeners' estimates of speech rate. Therefore, the current study seeks to suppress these potential effects by recruiting speakers who have minimal/no experience with the language used in the experimental stimuli to investigate syllable complexity in speech rate perception. If segment rate (again, syllable complexity) is relevant to speech rate perception, listeners should be sensitive to the manipulated segment rate of unfamiliar speech presented to them. They should judge stimuli with complex syllables as being faster than stimuli with simple syllables. The present findings support this hypothesis.

Methodology

Stimuli

The stimuli consisted of complete Arabic sentences. Arabic root-pattern morphology allowed us to manipulate and construct stimuli with two levels of syllable complexity, simple and complex. Set 1 of the stimuli consisted of twenty-four six-word sentences with simple syllable structure (*Mean* ≈ 1 C/syl) and a normal to slightly slow speech rate (*Mean* ≈ 3.9 syl/sec). Set 2 of the stimuli was identical to Set 1 in terms of the number of sentences and nearly identical in terms of the number of words per sentence. However, the segment rate in Set 2 was modestly increased, resulting in a relatively higher syllable complexity (*Mean* ≈ 1.75 C/syl) and hence the speech rate was relatively slower (*Mean* ≈ 3 syl/sec). An Arabic native recorded the sentences using the first recorded sentence as a model for the target speech rate of the remaining sentences. Then, before running the experiment, five Arabic natives were asked to judge the speech rate of the stimuli on a 1-7 scale (1 = extremely slow, 4 = neutral, 7 = extremely fast) in order to make certain that stimuli speech rate was just as intended (not extremely fast/slow). They generally found all sentences to have a neutral (4) or slightly slow (3) speech rate (*Mean* ≈ 3.80).

Procedure

Twenty-one male and female Japanese natives participated in a perception experiment. All participants had either no previous exposure to Arabic or, at most, knew a few simple words such /salaam/ "greetings". The age of the participants ranged from 21 to 39, and none of them reported any hearing loss or impairment. They were introduced to the experiment and provided with five trials for practice. The experimental stimuli were presented in an auditory form and in a randomized order. The participants were asked to make estimates about the speech rate, using the same scale used by the original Arabic judges.

Results and discussion

As illustrated in Figure 1, the participants generally judged the complex syllable stimuli as being faster (*Mode* = 3, *Mean* = 3.23) than the simple syllable stimuli (*Mode* = 2, *Mean* = 2.21). The participants judged almost 50% of the simple syllable stimuli as being slow and almost 50% of the complex syllable stimuli as being slightly fast. They also judged approximately 20% of the simple syllable stimuli and only about 2% of the complex syllable stimuli as being extremely slow, while they judged slightly more than 18% of the complex syllable stimuli and slightly less than 2% of the simple syllable stimuli as being neutral.

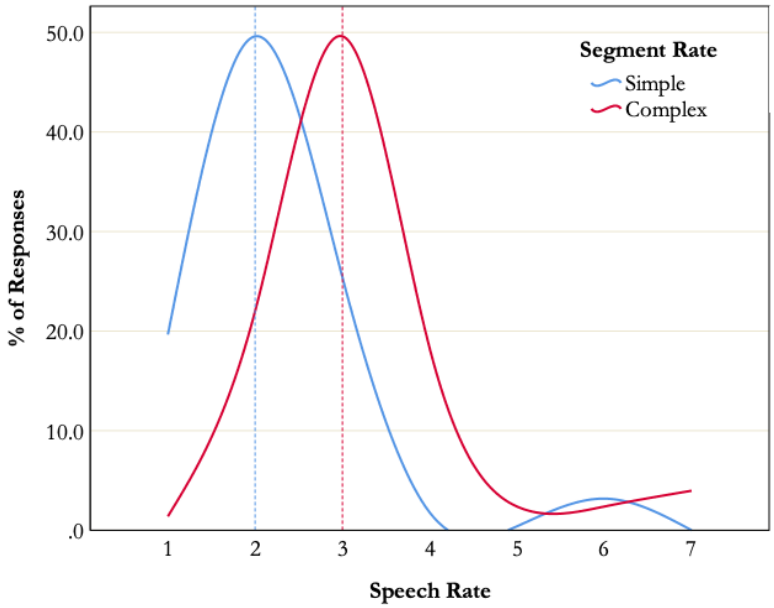


Figure 1. Proportional distribution of ratings of simple vs. complex stimuli

A repeated-measure regression model was performed with speech rate response as a predicted variable and segment rate/syllable complexity as a predictor variable. The model revealed that syllable complexity was an explanatory factor, Wald $\chi^2(1) = 147.99$, $p = 0.001$.

The results provide evidence that that segment rate does influence listeners' perception of speech rate. The difference in segment rate between the simple rate stimuli and the complex syllable stimuli is not that large (1 vs. 1.75 C/syl, respectively), but the this triggered a one-point increment in the ratings, both with respect to the means and modes of the data. Syllable rate (# of syl) cannot be the reason for the high ratings in the complex syllable stimuli because it was lower than that in the simple syllable stimuli. Nor can stimuli durations be a

cause, as they were roughly equal across stimuli at both levels of syllable complexity; in fact, the simple syllable stimuli were slightly shorter in duration.

Thus, the findings are in agreement with O'Dell and Nieminen (2019) who found that listeners are sensitive to segment rate, but are in (at least partial) disagreement with other findings which claim that segment rate does not play a crucial role in speech rate perception. Specifically, the pattern in the current findings is at odds with Plug and Smith's conclusion that there is "no evidence that listeners are using segment rate to influence their tempo judgements" (2018, p. 282). In contrast to Plug et al (2020), the influence of segment rate was not contingent on durational differences in the present study, as duration was controlled across stimuli to be as equivalent as possible. Indeed, if a slight difference in stimuli length were to have had any effect in the present study, it should have triggered higher ratings in the simple syllable stimuli (as in Plug et al, 2020) because they were marginally longer. However, this was not the case.

Conclusion

The findings show that differences in segment rate influence the perceived speech rate even if the difference in segment rate is not accompanied by a difference in syllable rate. This diverges from previous studies, possibly because the listeners were presented with stimuli from an unfamiliar language. Information load, syntactic parsing, and semantic processing did not bias the listeners' estimates; the listeners were found to judge the speech rate based on syllable rate and segment rate, with complex segment rates triggering higher ratings. Future studies may increase the complexity of the experiment design by adding conditions to vary syllable rate and segment rate in unfamiliar speech.

References

Aldholmi, Y., Park, H. 2019. Interactions between temporal acoustics and indexical information in speech rate perception. In Perspectives on Arabic Linguistics XXXI: Papers from the annual symposium on Arabic Linguistics, Norman, Oklahoma, 2017 (Vol. 8). John Benjamins Publishing Company.

O'Dell, M., Nieminen, T. 2019. Syllable rate, syllable complexity and speech tempo perception in Finnish. In Proceedings of the 19th International Congress of Phonetic Sciences 622-626, Melbourne, Australia.

Plug, L., Smith, R. 2018. Segments, syllables and speech tempo perception. In Proceedings of the 9th International Conference on Speech Prosody, 279-283, Poznan, Poland.

Plug, L., Lennon, R., Smith, R. 2020. Listeners' sensitivity to syllable complexity in speech tempo perception. In Proceedings of the 10th International Conference on Speech Prosody 2020. Tokyo, Japan.

Seifart, F., Strunk, J., Danielsen, S., Hartmann, et al. 2018. Nouns slow down speech across structurally and culturally diverse languages. Proceedings of the National Academy of Sciences, 115(22), 5720-5725.

Properties of nominal stress grammar in Greek

Vasiliki Apostolouda
Department of Linguistics, Aristotle University of Thessaloniki, Greece
https://doi.org/10.36505/ExLing-2020/11/0006/000421

Abstract

This study reports on the findings of a production experiment that aimed at exploring the distribution of stress patterns in the internal grammar of young and adult native speakers of Greek, with a special focus on nominal stress, as these are reflected on their decisions on assigning stress in novel inflected words. More specifically, we investigate whether morphological information at the right edge of the word, together with lexical frequency, are potential cues for stress position. We also examine differences in the grammars of young and adults speakers.

Keywords: lexical stress systems, nominal stress, pseudowords, grammar

Introduction

In this study we focus on both adult and younger speakers of Greek. The major objective is to investigate the impact of morphological information and of lexical frequency on their stress decisions. For this purpose, we designed and conducted an experimental task which involved the production of pseudowords with real inflection markers. In order to test the lexical frequency effects, we compared the experimental productions with the results of the quantitative research on the two lexical databases.

Greek is a lexical stress systems, that is, the position of stress is not predictable by phonological rules, but varies depending on the morphological structure of the word (Halle & Idsardi 1995; Alderete 1999; Revithiadou 1999). Nominal stress exhibits a broader variation, as stress patterns vary across the noun classes as well as within nouns of the same class. However, experimental studies on Greek (Apostolouda 2012; Revithiadou & Lengeris 2016) and other lexical systems (Lavitskaya 2014 for Russian) indicate that specific inflectional endings at the right edge of the word are potential cues for the stress position. In order to estimate the distribution of stress patterns over the various noun classes, we conducted a quantitative research on two lexical databases: (a) a Part of Speech (PoS) version[1] of Clean Corpus (CC) (Protopapas et al. 2012), (b) a PoS version of Reverse Dictionary of Modern Greek (RD) (Anastassiadis-Symeonides 2002). The results confirmed that specific noun classes are highly connected with specific stress patterns. A main goal of this study is to discover/unearth the extent in which this distribution is encoded in the stress grammar of native speakers. Several researchers have reported that lexical tendencies do not just leave a statistical imprint on the data but rather they are internalized in the native speakers' grammar (Zuraw 2000; Hayes & Londe

2006). Studies on Greek confirm these findings, as they reveal that native speakers tend to reproduce systematically the lexical stress patterns when they have to decide for the stress position of novel words with real class markers (Apostolouda 2012; Revithiadou & Lengeris 2016). Since previous research on the topic is solely based on adult speakers, it is important to also explore whether young speakers' decisions are affected by lexical frequency effects as well or whether their stress grammar exhibits early grammars' biases, (see the Emergence of the Unmarked, Prince & Smolensky 1993).

Methodology

Participants

The participants were divided into two groups based on their age. Group A consists of 47 adult native Greek speakers, 18-23 years old (mean age: 18.3), students of Aristotle University of Thessaloniki. Group B consists of 46 young native speakers, 7-8 years old (mean age: 7.7), students of 2nd and 3nd Grade of urban primary schools. The Group B participants were controlled for their performance at the standardized background test, Test of expressive vocabulary (Vogindroukas et al. 2009).

Stimuli

The experimental items are disyllabic and trisyllabic pseudonouns with real class markers (/-os/, /-as/, /-is/, /-a/, /-ifem/, /-o/, /-ineut/, /-(m)a/). The main goal was to construct pseudowords according to the phonotactics of Greek language by taking into consideration the degree of similarity with real words. For this purpose, we utilized specific research tools, as Clean Corpus and Num Tool (Protopapas et al. 2012). These tools can estimate the degree of similarity of each constructed pseudoword, providing quantitative measures for multiple variables (e.g., *Logmean bigram type & token frequency, Number of phonological neighbors, Phonological Levenshtein Distance*). Pseudowords with a mean score to each variable were considered to be the most suitable for the experiment. In a second round of evaluation, the stimuli were rated by native speakers of Greek. In a 5-scale Likert questionnaire, they had to decide for the degree of familiarity of each pseudoword. Pseudowords with a mean score to the degree of familiarity were finally selected as experimental items.

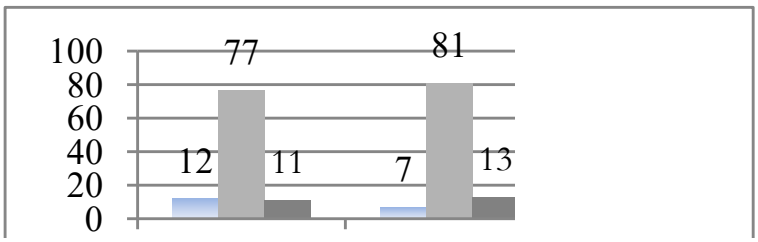
Procedure

The participants had to read out 128 disyllabic (2σ) and trisyllabic (3σ) pseudowords with no time restriction, deciding for the stress position. Pseudowords were presented in capital letters one at a time on Microsoft PowerPoint. In order to ensure that the young participants will complete the task, we embedded the stimulus in a playful context, where pseudowords were supposed to be the language of popular fictional characters. The same context

was utilized for the stimulus presentation to the adults. The utterances were recorded, transcribed and codified according to the experimental conditions.

Results

The statistical analysis revealed a significant difference between the frequency of penultimate stress and the frequency of the other stress patterns in 5932 items that collected from Group A and 5887 items from Group B (Fig. 1). Pseudonouns were primarily connected with penultimate stress, despite the type of inflection and the size of the word. However, the results between the two groups differentiate, if we focus to specific experimental conditions.



Group A: [U, PU] x2 = 2788.116, p < .001, [PU, APU] x2 = 2974,875, p < .001, [U, APU] x2 = 6,214, p = .013
Group B: [U, PU] x2 = 3684,173, p < .001, [PU, APU] x2 = 2863,773, p < .001, [U, APU] x2 = 122,510, p < .001

Figure 1. Stress distribution (%) in all items that collected from Group A and Group B (U=ultimate, PU=penultimate, APU=antepenultimate).

The participants' productions reveal two different tendencies. On the one hand, Group A's productions are largely affected by lexical frequency effects. For example, according to the lexicon, nouns in /-a/ are mainly associated with penultimate stress, while nouns in /-os/ and /-o/ are mostly associated with antepenultimate stress. In both conditions, Group A's productions consistently match the lexical frequencies. In contrast, Group B's decisions exhibit a clear tendency for penultimate stress, which is considered to be the unmarked pattern due to its trochaic nature ('σσ) (McCarthy & Prince 1994, see also Fikkert 1994 for child speech). They tend to overgeneralize penultimate in all conditions, regardless the type of inflection and the size of the word. The differences between the stress hierarchies that arise from both groups are clear in Table 1, especially if we focus on nouns in /-o/ and /-os/.

Table 1. Stress hierarchy for 3σ in /-a/, /-os/, /-o/.

Ending	Group A	Group B	CC	RD
/-a/	PU > APU > U	PU > APU > U	PU > U > APU	PU > APU > U
/-os/	APU > PU > U	PU >APU > U	APU > U > PU	APU > U > PU
/-o/	APU > PU > U	PU > APU > U	APU > PU > U	APU > PU > U

Conclusions

The findings of our research have revealed that lexicon tendencies and morphological information are activated in adult speakers' stress grammar. More specifically, adult speakers seem to systematically connect specific stress patterns with specific inflectional endings, replicating the effects of the Lexicon regarding stress. On the other hand, young speakers exhibit a strong bias for penultimate stress, the cross-linguistically unmarked stress pattern, across all the inflectional classes. This finding suggests that that lexicon frequencies and morphological information are not yet embedded in their grammar, which still shows traces of early grammars' biases, such as the Emergence of the Unmarked. Our findings indicate that the shaping of stress grammar is a dynamic process which is primarily controlled by linguistic factors but also crucially affected by lexical ones. Future research on the speech of (pre-) adolescent speakers may shed light on the way in which lexical factors interfere with linguistic ones.

Notes

1. The PoS version of Clean Corpus was developed by a research team of the Department of Linguistics at AUTh (Revithiadou et al. 2014; Apostolouda 2018).

References

Alderete, J. 1999. Morphologically Governed Accent in Optimality Theory. PhD dissertation, University of Massachusetts, Amherst.

Fikkert, Paula. 1994. On the Acquisition of Prosodic Structure. Ph.D. dissertation, HIL/ Holland Academic Graphics.

Hayes, B., Londe, S. J. 2006. Stochastic phonological knowledge: The case of Hungarian vowel harmony. Phonology 23, 59-104.

Lavitskaya, Y., Kabak, B. 2014. Phonological default in a lexical stress system. Lingua 150, 363- 385.

McCarthy, J. J., Prince, A. 1994a. The Emergence of the Unmarked: Optimality in prosodic morphology. North Eastern Linguistic Society 24, 333–379.

Prince, A. S., Smolensky, P. 1993. Optimality Theory: Constraint interaction in Generative Grammar. Unpublished manuscript, Rutgers University Center for Cognitive Science, New Brunswick, NJ.

Revithiadou, A., Lengeris, A. 2016. One or many? In search of the default stress in Greek. In Heinz, J., Goedemans, R., van der Hulst, H. (eds.) 2016, Dimensions of stress, 263-290. Cambridge, CUP.

Eliciting focus-sensitive *why*-questions in Japanese

Kodai Aramaki[1], Kanako Ikeda[2], Kyoko Yamakoshi[2], Tomohiro Fujii[1]
[1]Environment and Information Sciences, Yokohama National University, Japan
[2]Humanities and Sciences, Ochanomizu University, Japan
https://doi.org/10.36505/ExLing-2020/11/0007/000422

Abstract

The study argues that in focus-sensitive *why*-questions in Japanese, *why* must precede its focus associate. It is proposed that this word order restriction follows if the *why*-as-CP-modifier approach is applied to the Japanese construction under investigation. It also reports the results of the elicitation experiment conducted to experimentally confirm the word order restriction.

Keywords: focus, Japanese, question elicitation, *why*-question

Introduction

It has been observed that, unlike other *wh*-phrases, reason *wh*-adverbials such as *why* can be focus-sensitive (Bromberger 1992). *Why*-questions are affected by focus shift while other *wh*-questions such as *where*-questions aren't. (Focused items are in small caps.)

(1) a. Why did JOHN buy beer? -- Because he had his ID.
b. Why did John buy BEER (not wine)? -- #Because he had his ID.
(2) a. Where did JOHN find the key? -- In the basket.
b. Where did John find THE KEY? -- In the basket.

Kawamura (2007) observes that *because*-clauses in Japanese associate with focus, exhibiting a word order restriction. *Because*-clauses must precede foci.

(3) (*MARY-GA) [Peter-ga byoin-ni yobidasareta kara]
M-NOM P-NOM hospital-by called because
(MARY-GA) shiki-ni ressekishita.
M-NOM ceremony-to attended
'MARY attended the ceremony because Peter was called by the hospital.'

Now let us observe that this word order restriction on *because*-clauses carries over to their interrogative counterparts. That is, *naze* 'why', *nande* 'why', and *doosite* 'why', when associated with focus, must precede the focused item. (Following the literature, we call it their *focus associate*; see Jackendoff 1972, Rooth 1992, Erlewine 2014.) Although the judgements are somewhat subtle, a

contrast is found between the *why*-questions in (4a) and (4b). No such contrast is found between the *where*-questions in (5a, b).

(4) {a. Naze MAINITI / b. ??MAINITI naze} otokonoko-wa booru-o
why every day every day why boy-TOP ball-ACC
kick-Q
keru-no?
'Why does the boy kick a ball EVERY DAY?'

(5) {a. Dokode MAINITI / b. MAINITI dokode} otokonoko-wa booru-o
where every day every day where boy-TOP ball-ACC
keru-no?
kick-Q
'Where does the boy kick a ball EVERY DAY?'

The fact that *why*-focus order is required can be explained in the following manner. In their study of English *why*-stripping (e.g. *John bought beer. -- Why John?*) Yoshida et al. (2015) argue that focus-sensitive *why*-questions like (1a, b) are derived via base-generation of *why* in the matrix higher CP (Rizzi 2001; Ko 2005) and movement of the focus associate to the lower Spec,CP covertly.

(6) [$_{CP1}$ why C [$_{CP2}$ ____ C [$_{TP}$ … FOCUS …]]]

We propose that similarly in Japanese, a reason *wh*-phrase is base-generated in the higher CP and the focus movement *must* occur in LF. Then the focus-*why* order in (4b), where the focused item is apparently scrambled over *why*, can be ruled out as a violation of Proper Binding Condition (Fiengo 1977, Saito 1989). The focus associate would have to undergo lowering to the lower Spec,CP in order to get licensed, leaving an unbound trace. (The question of why scrambling of foci cannot be undone is left for future research.)

(7) [$_{CP1}$ FOCUS$_i$ [$_{CP1}$ why [$_{CP2}$ ____ [$_{TP}$ … t_i …] C] C]]]

As for the fact that (5a, b) do not differ in acceptability, we assume with Rizzi (1997, 2001) that *where* moves to Spec,CP2, which can be taken to mean that *mainiti* 'every day' is not a focused item that the wh-phrase associates with (Yoshida et al. 2015).

One question that arises is, to what extent are the judgements in (4-5) reliable? We believe the contrast in (4) is real but quite subtle and requires rich context. Furthermore, it is not clear at least to us exactly how pronunciation (i.e., prosody) interacts with word order in this construction. These considerations led us to use question elicitation rather than acceptability rating

to test this word order effect. This allows us to control for context and not to deal with prosody as an experimental variable in this preliminary study.

Experiment

We conducted a question elicitation experiment. The goal was to test the hypothesis that *wh* precedes the focus associate more often in reason questions than non-reason questions. Fifteen university students were tested individually. Two experimental conditions were examined in a within-subjects fashion: one condition attempted to prompt participants to ask *why*-questions and the other *where*-questions. In both conditions, the participants were instructed to hear stories, together with a puppet penguin, Pen-chan. Participants were told in advance that Pen-chan might not always pay attention and that they would be requested to ask him a certain question after each story just to see if he was focused. A sample story is shown in Table 1 below.

Table 1. A sample story, where each scene is accompanied with a picture shown on the computer screen.

Scene 1	Mother promises Ken that she will buy him a new soccer ball if he practices every day. He goes to a nearby playground. But it doesn't have a soccer goal. He wants to have one.
Scene 2	He decides to move to his school's playground, where a goal is available. He can practice a lot there.
Scene 3	He goes there to practice next day, too.
Scene 4	On the following day, too, he does the same. Mother gives him a new soccer ball.

After the story is told, the experimenter gives a prompt (e.g., (8)), and then the participant responds to it (e.g., (9)). Each participant experienced six critical trials and four fillers in a randomized order after two practice trials. Responses were coded into three categories: "wh-focus order," "focus-wh order," and "miscellaneous."

(8) *Experimenter's prompt*
Pen-chan-ni otokonoko-ga mainiti booru-o {a. *keru riyuu-o* / b. *basyo-o*} *kiitemite kudasai.*
'Could you ask Pen-chan the {a. reason why / b. place where} the boy kicks a ball every day?'

(9) *Sample wh-focus response to the Reason prompt*
Nee Pen-chan, otokonoko-wa nande mainiti booru-o keru-no?
hey Pen-chan boy-TOP why every day ball-ACC kick-Q
'Hello, Pen-chan. Why does the boy kick a ball every day?'

Results

The participants gave *wh*-focus responses 71.1 % of the time to Reason prompts (32/45) and 22.2% of the time to Place prompts (10/45). A one-way ANOVA revealed that the difference between the two conditions was highly significant ($F(1,14) = 19.05$, $p = 0.0006$). It showed that the participants know that *why*-questions are different from non-*why* questions in the way described by the theoretical analysis.

Discussion and conclusion

In this paper, we conducted an elicitation experiment to demonstrate a word order restriction on focus-sensitive *why*-questions in Japanese. The results revealed that the phenomenon exists and requires an explanation. The analysis we have proposed here constitutes one.

Acknowledgements

We thank Ken Hiraiwa, Riho Mochizuki, and Masaya Yoshida for their helpful comments on earlier versions of this work.

References

Bromberger, S. 1992. On What We Know We Don't Know: Explanation, Theory, Linguistics, and How Questions Shape Them. Chicago, University of Chicago Press.

Erlewine, M. Y. 2014. Movement out of Focus. Doctoral dissertation, MIT.

Fiengo, R. 1977. On trace theory. Linguistic Inquiry 8, 35-61.

Jackendoff, R. 1972. Semantic Interpretation in Generative Grammar. Cambridge, MIT Press.

Kawamura, T. 2007. Some Interactions of Focus and Focus Sensitive Elements. Doctoral dissertation, Stony Brook University.

Ko, H. 2005. Syntax of *why*-in-situ: Merge into [Spec, CP] in the overt syntax. Natural Language and Linguistics Theory 23, 867–916.

Rizzi, L. 1997. The fine structure of the left periphery. In Haegeman, L. (ed.) Elements of Grammar, 281–337. Dordrecht, Kluwer.

Rizzi, L. 2001. On the position "Int(errogative)" in the left periphery of the clause. In Cinque, G., and Giampaolo, S. (eds.) Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi, 267–286. Amsterdam, Elsevier.

Rooth, M. 1992. A theory of focus interpretation. Natural Language Semantics 1, 75-116.

Saito, M. 1989. Scrambling as semantically vacuous A'-Movement. In Baltin, M., Kroch, A. (eds.) Alternative Conceptions of Phrase Structure. Chicago, University of Chicago Press.

Yoshida, M., Nakao, C., Ortega-Santos, I. 2015. The syntax of *why*-stripping, Natural Language & Linguistic Theory 33, 323-370.

Comprehension of verb directionality in LIS and LSF

Valentina Aristodemo[1], Beatrice Giustolisi[2], Carlo Cecchetto[2,3], Caterina Donati[1]

[1]Laboratoire Linguistique Formelle, Université de Paris, France
[2]Department of Psychology, University of Milan – Bicocca, Italy
[3]UMR 7023 Structures Formelles du Langage, Université Paris 8, France
https://doi.org/10.36505/ExLing-2020/11/0008/000423

Abstract

The present work reports the results of a comprehension task on verb directionality in Italian Sign Language (LIS) and French Sign Language (LSF) considering native and non-native signers. Our goals were to study age of first language exposure effects on the comprehension of verb agreement in LIS and LSF, to verify whether a significant difference between forward and backward directionality was found, and see if our results may provide insight about the nature (gestural vs. linguistic) of verb directionality in sign languages. In both languages we found that the ability to comprehend verb agreement is affected in non-native signers. This indicates that delayed first language exposure has long lasting effects in adulthood. We argue that our results support analyses of verb agreement as a fully grammatical phenomenon.

Keywords: agreement, directional verbs, sign language, age of exposure, Italian, French

Introduction

In sign languages (SLs), directional verbs move from the position in space associated to an argument towards the position associated to another argument. These verbs have been called 'agreeing verbs' by Padden (1988).

The interpretation to be given to of verb directionality in SLs has triggered much discussion and controversy among linguists. While some scholars use the term 'agreeing verbs' and insist that directionality is a fully grammatical phenomenon (Padden 1988, Pfau et al. 2018, a.o.), other scholars (e.g. Liddell 1995) deny that verb directionality can be assimilated to agreement and use a different terminology for these verbs, which are sometime called 'indicating verbs' or 'directional verbs' (Pfau et al. 2018 for an assessment of this literature) insisting on their gestural nature. For concreteness, in this paper we use the term 'agreeing verbs' to refer to these verbs.

Agreeing verbs can be further categorized into forwards and backwards verbs. Forwards verbs incorporate a movement going from the subject position to the object position, while backwards verbs show the inverse pattern, i.e. a movement from the object position to the subject position.

Behavioral studies have shown that comprehension of verb directionality is affected by age of first language exposure (AoE) in American Sign Language

(ASL) and British Sign Language (BSL) (Emmorey et al. 1995, Cormier et al. 2012, a.o.). AoE is a crucial variable to consider when assessing SLs comprehension: in fact, only a minority of deaf children are native signers, namely they acquire a SL from birth having at least a deaf signing parent. The majority of deaf children have hearing speaking parents, and if they are exposed to a SL, this typically occurs in kindergarten, or later.

In the present work, we investigated the comprehension of verb directionality in LIS and LSF. We collected data from 3 groups of signers: i) native signers (exposed to a SL from birth and with at least one signing parent), ii) early signers (AoE between 2 and 6), iii) late signers (AoE between 6 and 15). Our goals were: i) to evaluate whether AoE has an impact on directionality comprehension in LIS and LSF; iii) to verify if there is any difference in comprehension between forwards and backwards verbs; iii) to investigate whether effects of AoE can contribute to the debate about the nature of verb directionality.

The present study

Participants saw a brief non-linguistic clip showing three characters interacting, followed by a LIS/LSF sentence containing an agreeing verb, either forwards or backwards, and they had to judge whether the sentence matched the situation described in the clip. The sentence could either correctly describe the situation (match condition), or describe the situation attributing different thematic roles to the characters or displaying wrong argument selection (mismatch condition. Sentences were always signed by character A, who was therefore the grammatical first person, to character B (the grammatical second person). In the example, the matched sentence was "I yelled at you", whereas a mismatched sentence could be either "You yelled at me" (thematic roles inversion) or "I yelled at C" (wrong argument selection). Control sentences, which were clearly wrong, were added (like "I kissed you" in relation to the aforementioned example).

Materials

In LIS, the task included 18 situations and 4 sentences for each condition (1 correct, 2 with mismatched agreement, 1 control), for a total of 72 sentences. 12 forwards verbs and 6 backwards verbs were used. Participants saw all 72 sentences divided in two blocks.

In LSF, the task included 24 situations and 2 sentences for each situation (1 correct and 1 with mismatched agreement) and 20 control sentences, for a total of 68 sentences. 18 forwards verbs and 3 backwards verbs repeated twice were used. Participants saw the 68 sentences divided in two blocks, which were administered in two separate days.

In both tasks, items were pseudorandomized so that the same situation could not be repeated twice in a row.

Participants

Forty-one Deaf LIS signers participated. Three participants were excluded because their score was below 75% accuracy in control sentences. The final sample consisted of 38 signers (mean age=46 yrs, sd=9 yrs): 14 native, 13 early and 11 late.

As for LSF, 49 Deaf signers participated. Five participants were excluded because they were exposed to sign language when they were older than 15. Four participants were excluded because their score was below 75% accuracy in control sentences. The final sample consisted of 40 signers (mean age=37 yrs, sd=9 yrs): 14 native, 14 early and 12 late.

Results

In LIS, no clear pattern emerged between forwards and backwards verbs, and native signers seemed to perform better than early and late signers in the mismatch condition only. In LSF, performance looked worse for backwards verbs in the match condition, and in general slightly better for native signers compared to the other groups.

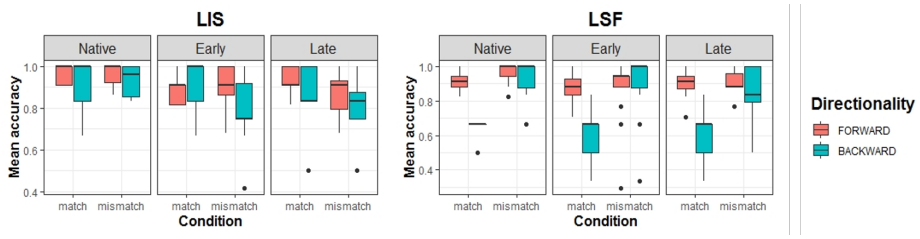


Figure 2. LIS and LSF results considering group, condition and verb directionality.

Analysis was performed separately for each SL, and for matching/mismatching sentences. We used generalized linear mixed models with random intercepts for subjects and items, and by-subject random slopes for the effect of directionality and by-item random slopes for the effect of group. Fixed factors were directionality, group, their interaction and participants' age. In LIS, native signers were more accurate in rejecting mismatching stimuli than early (β=-1.83, SE=0.54, z=-3.41, $p<0.001$) and late signers (β=-1.94, SE=0.54, z=-3.66, $p<0.001$). No significant difference was found comparing late and early signers. Verb directionality did not influence accuracy and no significant result was found in the match condition analysis. In LSF, native signers were more accurate than early and late signers in accepting matching and rejecting mismatched sentences (match – early: β=-1.45, SE=0.48, z=-3.05, p=0.002 ; late: β=-1.11, SE=0.47, z=-2.37, p=0.018; mismatch – early: β=-1.71, SE=0.62, z=-2.78, p=0.005 ; late: β=-1.10, SE=0.58, z=-1.88, p=0.06). In accepting matching sentences, verb directionality had a marginal impact (p=0.08).

Discussion and conclusion

In both languages, we found that natives outperformed non-native signers. This result shows that the capacity of interpreting directionality in LIS and LSF is affected when access to language is delayed, coherently with previous results in other SLs. To us, this result is consistent with those approaches that take directionality to be a fully linguistic phenomenon, and can be interpreted as another evidence for the existence of a sensible period for language acquisition. As nothing similar to sensitive periods has been established (yet) for gestural systems, our findings do not support approaches that take directionality to be gestural. Moreover, our results highlight that delayed first language exposure has long lasting effects in adulthood. This finding, replicated across languages and linguistic structures, gives clear information for health-policy makers dealing with prelingual hearing loss.

Only in LSF we found a marginal effect of inverse directionality (backwards verbs were slightly more difficult than forwards verbs). This cross linguistic difference needs further investigation.

Acknowledgements

We are grateful to all participants who took part in the study. Many thanks to: Anna Folchi, Rosella Ottolini, Mirko Pasquotto, Mauro Mottinelli, Gabriele Caia, Antonino D'Urso for their contribution as LIS consultants; Valerie Jendoubi for her contribution as LSF consultant. This research was supported by the SIGN-HUB project (European Union's Horizon 2020 research and innovation program, Grant Agreement N 693349).

References

Cormier, K., Schembri, A., Vinson, D., Orfanidou, E. 2012. First language acquisition differs from second language acquisition in prelingually deaf signers: Evidence from sensitivity to grammaticality judgement in British Sign Language. Cognition 124, 50-65.

Emmorey, K., Bellugi, U., Friederici, A., Horn, P. 1995. Effects of age of acquisition on grammatical sensitivity: Evidence. Applied Psycholinguistics 16, 1-23.

Liddell, S. 1995. Real, surrogate, and token space: Grammatical consequences in ASL. In Emmorey, K., Reilly, L. (eds.) 1995, Language, gesture, and space, 19-41. Hillsdale NJ, Erlbaum.

Padden, C. A. (1998). The ASL lexicon. Sign language & linguistics 1, 39-60.

Pfau, R., Salzmann, M., Steinbach, M. (2018). The syntax of sign language agreement: Common ingredients, but unusual recipe. Glossa: a journal of general linguistics 3, 107.

Complex syntax intervention for Developmental Language Impairment

Maria Athanatou, Elena Theodorou
Department of Rehabilitation Sciences, Cyprus University of technology, Cyprus
https://doi.org/10.36505/ExLing-2020/11/0009/000424

Abstract

Intervention for children with Developmental Language Disorder appears to be beneficial and contributes to sustainable linguistic gains. This paper reports on a pilot intervention study carried out in Cyprus that examined the efficacy of language treatment targeting complex syntactic structures. Language skills of a nine-year old girl with DLD are described at two time points, before and after intervention. The child received therapy sessions based on MetaTaal therapy, and relative clauses were the targeted syntactic structures. Post-intervention measurements showed marginal improvement in relative clauses production and comprehension. Improvements observed in Complex Sentence Repetition Task and this might imply that the grammatical structures have emerged.

Keywords: DLD, relative clauses, metalinguistic approach, syntax

Introduction

The term Developmental Language Disorder (DLD) was proposed (Bishop, et al., 2016) to describe the language difficulties and functional limitations that may coexist or appear in children with oral language deficits such as poor school performance and difficulty in forming social relationships. According to the international literature, 7% of all typically developing children experience DLD (e.g. Tomblin, Records, & Zhang, 1996; Norbury et al., 2016).

Focusing on the language characteristics, DLD causes problems in various aspects of language such as phonology, morphology, word finding, syntax and pragmatics in comprehension and/or production (see Leonard, 2014 for a review). Although it becomes clear that DLD is characterized by great heterogeneity (Friedman, 2007; Leonard, 2017), problems with morphology and syntax affect the majority of children who experience the disorder (e.g. Fey et al., 2004; Leonard, 2014; Nippold et al., 2009). Most common difficulties are the production and comprehension of complex syntactic structures such as relative clauses (RCs). Children with DLD seem to struggle with the production and comprehension of RCs due to the many different types of the particular clauses, gender and subject/object agreement, different positioning of verbs in clauses among others (e.g. Novogrodsky & Friedman, 2006).

Speech and language intervention research highlights that it can bring positive outcomes not only in relation to the language abilities of the child but also in other domains such as school performance and social involvement etc.

(e.g. Lousada, et al., 2016; Hegazi, et al., 2020). Balthazar and her colleagues (2020) suggested that clinicians should select specific syntactic targets based on individualized analysis and implement implicit learning principles and explicit metalinguistic instruction. This can be attained by planned selection and organization of stimulus materials while activities using multiple modalities occur.

Recent studies have suggested that metalinguistic approaches which target complex syntax, such as Shape Coding and Colourful Semantics could be beneficial especially for older children with DLD (Zwitserlood et al., 2015). In 2015, Zwitserlood et al. (2015), used the metalinguistic and multimodal intervention programme called 'MetaTaal' using 'Lego Bricks' and showed that RCs production was improved significantly.

The aims of our study were to examine the effect of the metalinguistic treatment targeting complex syntactic structures on language performance and on the production and comprehension of RCs.

Methodology

In this study a 9;7 y.o girl, E., had been experiencing difficulties with oral expression and understanding, and had a profile consistent to the diagnosis of DLD. It is noted that E. had received speech therapy sessions in the past but she still experienced language struggles. During the first phase of this experimental study, three assessment sessions took place, in which language tests, standardized and experimental, were administrated. The tests used for the purposes of this study were: The Bus Story Test (Renfrew, 1995), Sentence Repetition Task (Theodorou, et al., 2017) and RCs comprehension and production tasks (Theodorou & Grohmann, 2012), in order to note a baseline, and then compare those measurements after the intervention period. In order to confirm nonverbal intelligence scores within the average range Raven's Progressive Matrices test was administrated. A summary table with the results before and after the intervention programme is given below.

During the second phase, 8 intervention sessions that targeted production and comprehension of RCs were carried out. The intervention programme was an adaptation of the 'MetaTaal' approach (Zwitserlood, et al., 2015). We modified the programme with the replacement of lego bricks by coloured cords that are used as stimulus material which represented words of various categories like subject, relative pronoun etc.. Each intervention session was devoted to explicit metalinguistic instruction and guided practice by using the proposed activities. Firstly, the child practiced with exercises aimed at the identification of different clause types and conjunctions in spoken language. Then, she started constructing subject and object RCs using the coloured cards. The next step was to create full sentences and connect them by using the cards. Then, she was asked to connect the two sentences by erasing the subject of the 2nd clause. After that, the participant had to create a simple Subject-Verb-

Object sentence and use the bridge card (that represented the relative pronoun *that* 'pu'), in order to create a subject or object relative clause. The next step was to create right branching subject and object RCs, were the small bridge was included which it represented coma in written speech and pause in verbal speech. Finally, the participant had to create center-embedded subject and object RCs also by using the guide. Important to note that, in all stages of the intervention programme different pictures were used, all the instructions were given orally and no homework or written task was given. Further, all complex sentences used in the treatment were drawn from a pool of sentences developed for the purposes of the project by the authors.

Results

Table 1. Pre and post intervention scores on language tasks

Language Tests	Pre intervention programme measurements	Post intervention programme measurements
The Bus Story Test	Mean Length of Utterance-Words: 10,6	Mean Length of Utterance-Words: 9,2
	Subordinated Clauses: 12	Subordinated Clauses: 13
Complex Sentence Repetition Task	9/24	18/24
RCs production task	9/20	10/20
RCs comprehension task	17/32	16/32

Starting with the Bust Story Test that assesses narrative abilities, E.'s Mean Length of Utterances (MLU) was 10,6 words and she produced 12 subordinated clauses before the intervention programme. After the intervention programme E.'s MLU was 9,2 words and produced 13 subordinated sentences. With regards to the Complex Sentence Repetition Task, improvement was shown since she repeated correctly 9 out of 24 sentences before the treatment period and 18 after. Focusing on the the RCs comprehension and production tasks, comparing the number of correct answers no improvement is exhibited. As it is depicted in the table 1, E. comprehended 17 out of 32 RCs before the intervention programme and 16 after. The reduction in the number of errors concerns subject RCs. In addition, regarding the RCs production task, 9 out of 20 targeted sentences were produced before the intervention programme while 10 correct productions were noted after. It is important to note that the wrong answers in the second testing were only about object RCs. Specifically, a systematicity in the type of errors (reversal error) made in the comprehension task observed after the intervention programme, for the object RCs. As for the production task, the range of errors reduced, where high percentages of gap

filling and change of word order were noted. These two types of errors produced structures that approach the correct and targeted structure.

Discussion

The aim of this pilot study was to investigate the effect of the metalinguistic treatment targeting complex syntactic structures on language performance and on the production and comprehension of RCs.

Results suggest that the metalinguistic and multimodal intervention used here can produce improvements in Complex Sentence Repetition task. This might indicate that complex grammatical structures have emerged although the child cannot generate them simultaneously which is in line with the assumption that sentence repetition is an indicator of child's language ability (e.g. Marinis & Armon-Lotem, 2015; Polišenská et al., 2015).

As for the length of utterances, which is considered an indicator of the complex syntactic abilities, no improvement was observed. However, more analysis is needed that will examine potential differences in terms of the types of syntactic structures that E. used.

In conclusion, it is shown that the treatment approach used can be effective for children with DLD but more research is needed in order to maximize gains and enhance the impact to language growth.

References

Bishop, D.V., Snowling, M.J., Thompson, P.A., Greenhalgh, T., Catalise Consortium. 2016. CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. PLOS one, 11(7), e0158753.

Leonard, L.B. 2014. Children with specific language impairment. MIT press.

Theodorou, E., Kambanaros, M., Grohmann, K.K. 2017. Sentence repetition as a tool for screening morphosyntactic abilities of bilectal children with SLI. Frontiers in Psychology, 8(DEC). https://doi.org/10.3389/fpsyg.2017.02104

Zwitserlood, R., Wijnen, F., Van Weerdenburg, M., Verhoeven, L. 2015. "MetaTaal": Enhancing complex syntax in children with specific language impairment-a metalinguistic and multimodal approach. International Journal of Language and Communication Disorders, 50(3), 273–297. https://doi.org/10.1111/1460-6984.12131

Disambiguating cues of disjunctive questions

Mohammad Ali. S. Bani Younes, Sam Hellmuth
Language and Linguistic Science, University of York, UK
https://doi.org/10.36505/ExLing-2020/11/0010/000425

Abstract

This paper aims to find out the similarities and differences between Jordanian (JA), Egyptian (EA), and Kuwaiti (KA) Arabic in which cues disambiguate alternative questions (altqs) and disjunctive yes-no questions (dynqs): intonation contour and choice of disjunctive element (DE). A perception study was run in the three dialects, replicating Pruitt & Roelofsen's (2013) perception study on English. Mixed-effects logistic regression was used to explore the results which revealed all dialects show a main effect of both intonation and DE choice; a rise contour and use of *ʔaw* significantly increased the likelihood of dynq responses. The effect of intonation was larger than that of DE choice in all dialects. The differences between the dialects lay in the relative strength of the DE coefficients.

Keywords: Alternative questions, disjunctive yes-no questions, intonation, prosody.

Introduction

A question like *Do you want coffee or tea?* can be perceived differently in English: either as an altq (possible answer: Tea) or as a dynq (possible answer: No). The two readings of the string-identical utterance are perceptually disambiguated by prosodic cues: with a fall, the utterance is perceived as an altq, but with a rise, as a dynq (Pruitt, Roelofsen 2013, Dayal 2016). They also claim that placing accents on both disjuncts 'X' and 'Y' (in the 'X or Y') favours an altq reading. One reason why these questions may be disambiguated by prosody in English is that there is only one DE *or* which is used in both question types. Modern Standard Arabic (MSA) has two DEs: *ʔam* in altqs and *ʔaw* in dynqs (Fakih, 2012). Colloquial Arabic dialects also display two DEs - *willa* and *ʔaw* - which vary in their mapping to English *or*. To replicate Pruitt & Roelofsen's study in Arabic, therefore, one must first decide which DE should be used in such a perception study.

A review of corpus data and prior literature suggests that some Arabic dialects might be English-like in employing one DE in both altqs and dynqs (e.g. EA) while other dialects might be MSA-like in using two DEs, one for altqs and one for dynqs (e.g. KA). Previous research (Bani Younes, Hellmuth 2020) has found that intonation (rise) and DE choice (*ʔaw*) both favour a dynq reading in JA though both DEs are accepted in both question types, placing JA in the middle of the two preferences between English and MSA. Having established the disambiguating cues, how similar or different are the three dialects in these cues? And what is the relative contribution of each of the two

cues in the disambiguation? The present study, thus, adapts Pruitt & Roelofsen's design for JA, EA, and KA, using the two cues: intonation (rise/fall) vs. DE choice (*willa*/*ʔaw*). Figure 1 below shows the same utterance produced with both contour shapes and DEs.

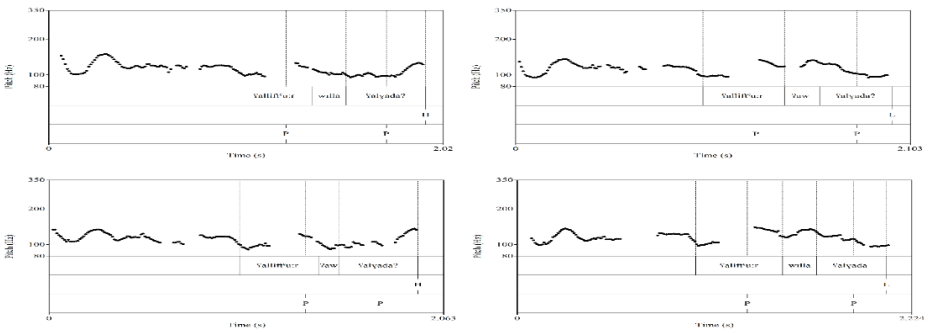


Figure 1. Contour shape vs. DE choice of *lljom ʕazmatak ʔæia ʕalifᶜu:r ʔaw/willa ʕalyada*? 'Did Aya invite you to breakfast or lunch (to have breakfast or lunch)?'

Methodology

There were 74 JA, 52 EA, and 70 KA listeners (mean age 36; range 18-53). 24 lexically-distinct target sentences and 36 fillers were recorded. The task was a two-alternative forced choice between an altq and a dynq paraphrase of the question they heard. Each target utterance was recorded in four conditions (see Figure 1 above), yielding 96 target tokens; 24 unique tokens were presented in each of four blocks in Qualtrics, in a Latin-Square design. The four conditions in each block are: 6 *ʔaw*+rise (2r), 6 *ʔaw*+fall (2f), 6 *willa*+rise (wr), and 6 *willa*+fall (wf) tokens. Participants were sent the experiment link and were asked to listen to the recordings via their own laptop or mobile phone in a quiet room, free from noise, using ear/headphones. After they clicked on the link, participants were randomly assigned to one of the four blocks containing 60 tokens (24 unique target utterances and 36 fillers). Two paraphrases appeared below each recording, so participants ticked one of them. Paraphrases were randomly ordered for each participant within each trial, and trials were also randomly ordered. After listening to each token, they selected the paraphrase that they thought was the best equivalent to what they had heard.

Results and discussion

Table 1 presents the results of the three separate GLMM models (one per dialect: glmer(resp_numeric~intonation*DE + gender + age +Education + (1 + intonation | listener) + (1 | stimulus), data = data2, family = binomial, control = glmerControl(optimizer = "bobyqa"))). Dynqs were coded as 1 while altqs as 0. Figure 2 shows the spread of values across participants for counts of tokens interpreted as dynqs, by dialect.

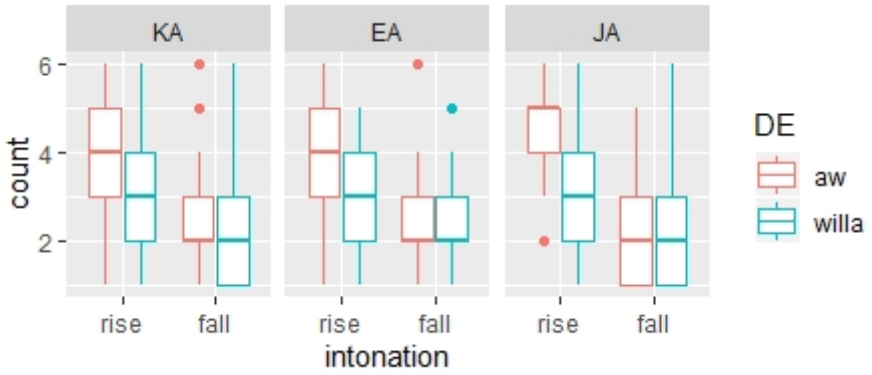


Figure 2. Median and interquartile range for count of dynq responses across participants for each dialect.

	Fixed effects	Estimate	SE	z value	p-value	
JA	Intercept	-0.53423	0.52680	-1.014	0.311	
	intonation1	0.95554	0.11590	8.244	2e-16	***
	DE1	0.34167	0.08524	4.009	6.11e-05	***
	intonation1:DE1	0.12901	0.08517	1.515	0.130	
EA	Intercept	-0.2955836	0.4127871	-0.716	0.4739	
	intonation1	0.4677226	0.1001999	4.668	3.04e-06	***
	DE1	0.3723088	0.0872576	4.267	1.98e-05	***
	intonation1:DE1	0.1470266	0.0871239	1.688	0.0915	.
KA	Intercept	-0.225396	0.436883	-0.516	0.6059	
	intonation1	0.602045	0.089545	6.723	1.78e-11	***
	DE1	0.147364	0.066593	2.213	0.0269	*
	intonation1:DE1	0.065525	0.066591	0.984	0.3251	

Table 1. Estimates of coefficients in JA, EA, and KA.

The control variables (gender, age, education) shown in the model syntax above were not included in the table as they did not reach significance. There are main effects of intonation and of DE choice in each dialect, suggesting that both cues are important in all three dialects. However, the effect size of DE choice compared with intonation is subtly different across the three dialects. JA, EA, and KA are thus similar in that both intonation and DE choice significantly influence the likelihood of dynq responses. Likewise, it seems that intonation was more important than DE choice in increasing the likelihood of dynq responses in the three dialects. This matches the results of prior studies on the role of intonation in characterising the realisation of altqs and yes-no questions (ynqs) in JA, EA, and KA (e.g., Al-Khalifa 1984, Rammuny 1989, Al Amayreh 1991, Alharbi 1991, Hellmuth 2018, Winans, 2019).

The findings also corroborate the fact that the contour shape of dynqs is the same as that of ynqs. In other words, the fact that tokens with a rise

significantly favoured dynq responses, over altq responses, is also in line with Winans' (2019) claim that *ʔaw*-dynqs have the same contour as ynqs in EA; she explained this similarity by the fact that both are ynqs. This similarity in the contour of dynqs and ynqs might explain listeners' tendency to interpret rise-tokens more as dynqs in JA, EA, and KA. For this reason, interpreting tokens with a rise as dynqs is unsurprising.

The results show slight differences between the dialects. The coefficient estimates for intonation and DE in EA are similar in size, suggesting that the relative importance of DE choice in EA was the strongest across the three dialects, although intonation still plays the main role in the disambiguation. JA and KA were similar in having DE coefficient values which are much lower than those of intonation, but the effect of DE choice in KA was still smaller than that of DE choice in JA, making KA different from JA in this respect, too.

In conclusion, the findings are both similar to and different from Pruitt and Roelofsen's (2013) study on English, taking into account the idiosyncratic cues of each language. Pruitt and Roelofsen's findings suggest that two cues (intonation vs. accent distribution), both prosodic, were significant in disambiguating altqs and dynqs. They also show that intonation played the primary role while accent distribution played a supporting role. This paper adds a lexical cue (DE choice) and finds that both cues in Arabic, which are prosodic and lexical, are also needed, though the primary role is again associated with intonation and the secondary role with DE choice.

References

Al Amayreh, M.M.K. 1991. A study of the basic intonational patterns in Standard and Jordanian Arabic (Master's thesis). University of Jordan. Amman, Jordan.

Alharbi, L.M. 1991. Formal analysis of intonation: The case of the Kuwaiti Dialect of Arabic. Ph.D. Dissertation. Heriot-Watt University, UK.

Al-Khalifa, F.E. 1984. An intonational study of Kuwaiti Bedouin Arabic. Ph.D. Dissertation. University of London, UK.

Bani Younes, M., Hellmuth, S. 2020. The role of prosody in disambiguation of disjunctive questions in Jordanian Arabic. Paper presented at Arabic Linguistics Forum Conference, University of Leeds.

Dayal, V. 2016. Questions. Oxford University Press.

Fakih, A. H. 2012. The syntax of questions in Modern Standard Arabic: A minimalist perspective. Germany: LAMBERT Academic Publishing.

Hellmuth, S. 2018. Variation in polar interrogative contours within and between Arabic dialects. Proceedings of Speech Prosody. Poznań, Poland, 13-16 June.

Pruitt, K., Roelofsen, F. 2013. The interpretation of prosody in disjunctive questions. Linguistic Inquiry, 44(4), 632-650.

Rammuny, R. 1989. Instrumental and auditory analysis of Colloquial Jordanian Arabic, an update. Journal of Arabic Linguistics (20) 23-42.

Winans, L. 2019. Disjunction and alternatives in Egyptian Arabic In K. von Heusinger et al. (Eds.), Questions in Discourse (pp. 237–285). Leiden, The Netherlands: Brill.

On only-pragmatically driven intonation change

Marco Barone
Department of languages and literature, Federal University of Pernambuco, Brazil
https://doi.org/10.36505/ExLing-2020/11/0011/000426

Abstract

The intonation system of the Italian variety of Pescara was documented and two sentence types (neutral polar questions and contrastive focus statements) were found to exhibit the same two pitch accents as allophonic variants by the old speakers. However, moving on the new generation, the variations of the two sentence types shows different evolutions: both variants are used, remaining distinct, for contrastive focus, whereas they mainly fuse into a "midway" pattern, when used for questions. The asymmetry can only be ascribed to the pragmatics and not to the phonetic forms of the patterns, as these were originally equal across the two sentence types. This suggests that polar questions are more kin to phonetic convergence than contrastive statements.

Keywords: intonation phonology, intonation change, discourse completion task, post-lexical meanings, phonetic convergence

Introduction

Intonation phonology studies the conventional associations, in a given intonation language or variety, between surface phonetic forms (contours and their smaller units) and post-lexical meanings, also referred to as "sentence types" (specifying modality and pragmatic function, including its nuances). Roughly speaking, it can be though of as a type of "phono-pragmatics" (as opposed to the classical "phonosemantics"): a correspondence, not always one-to-one, between a phonetic level (signifier) and a pragmatic level (signified).

In the intonational sign, just like in standard phonology, the signifier is subject to variation across languages, whereas meanings are mainly conceived as a fixed, universal layer (although cross-cultural differences in their categorization and their reanalysis in time cannot be excluded). This is why studies on intonation variation and change usually work one sentence type at a time and compare the contours used for it. Since these studies focus on one or a few sentence types only, they usually build their corpus by eliciting only those, and analyzing the variation of contours (or smaller patterns) associated to them. However, when an asymmetry between two signs is found, it is hard to assess whether it is due to an asymmetry of the signifiers, or of the meanings, unless either the signifiers or the meanings coincide across the two signs.

Intonation change studies mainly amount to contact-induced change. Among them, some explore variation across age groups (Heffernan 2006, Alvord 2010). As change is not uniform across speakers, it can also be detected by the variation it leaves behind, especially in contact situations, where control groups

help detect which are, for a given group, the native (conservative) variants and those of exposure (innovative) and thus the direction of change.

Recent studies have seen change as the long-term outcome of the actual interaction of speakers of mutually intelligible systems (varieties of the same language), rather than occurring in the mind of bilingual speakers, supporting this way hypotheses such as indirect transfer and language accommodation theory: convergence of speakers in daily interaction leads to the surface forms of the systems becoming more similar to one another, unconsciously and gradually (phonetic convergence). This may lead to interlanguages with phonetically intermediate patterns, which in segmental phonology Chambers & Trudgill (1980) refer to as "fudged lects", as opposed to "mixing lects", which contain patterns from both donor languages but no further ones.

Studies on intonation contact, transfer and accommodation have found the intonation contours of yes-no questions more prone to change towards nonnative contours than those of statements. O'Rourke (2005) observes that Cuzco speakers maintained their conservative L1-Quechua contour in their L2 Spanish for declaratives, but they acquired the L2-Spanish pattern for polar questions. Romera & Elordieta (2013) find that the contours used by a Catalan-dominant and a Spanish-dominant groups of Majorcan speakers for statements and those used for questions evolved differently: the two groups' original patterns would phonetically converge to an intermediate form only in the case of questions. Dealing with different allophonic distributions, the authors did not ascribe the asymmetric behavior to the pragmatics of questions vs statements and rather posit a role of the difference in salience of the contours' shape, that is, a (language specific) characteristic of the surface form. Another study (Troncoso-Ruiz & Elordieta 2017) finds that Andalusian speakers reallocated in Asturias accommodate their behavior to the local intonation more frequently for questions than for statements. Again, as the contours of the two varieties are more distinct for questions than for statements, the authors ascribe the asymmetry to salience of the signifier.

Methods

Our corpus is part of a survey on Pescara Italian intonation, made within the Interactive Atlas of Romance Intonation (Prieto, Borràs-Comes & Roseano (Coords.) 2010-2014), using the Discourse Completion Task (Blum-Kulka et al. 1989, Vanrell et al. 2018), a role-play interview aimed at eliciting different speech acts, by presenting imaginary contexts participants are supposed to react to by uttering, for each one, a specific target sentence type. The task was adapted to Italian by Barbara Gili Fivela (Gili Fivela et al. 2015) and contains 57 situations, targeting 35 sentence types. The documentation of Pescara Italian intonation included 6 old speakers (63-68 y.o.) and 8 young speakers (22-37 y.o.), for it was aimed at capturing information on the influence of the substratum language the elder speakers are more proficient in.

Data analysis and results

Both for contrastive statements and information-seeking yes-no questions, old speakers of Pescara Italian show the **same allophonic variation** of pitch accents, with 2 main pitch accents (Figure 1): A) a "pre-tonic peak" fall, transferred from the local language, Pescarese, with a high peak aligned at the end of the pre-tonic syllable, and B) a "tonic peak" late fall with a peak fully inside the tonic vowel, influenced from Rome Italian, together with a minority of intermediate patterns, being either **overlapping patterns** ("double" peak (C1) or plateau (C2)) or a **midway pattern**, a single "midway peak" (D) usually aligned between the onset and the nucleus of the tonic syllable.

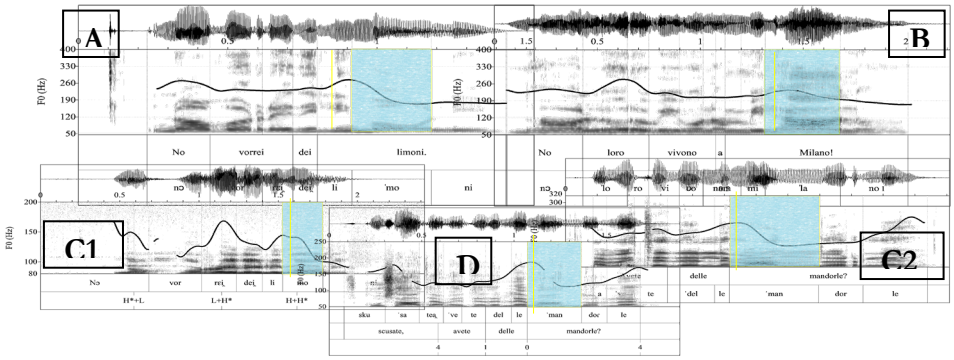
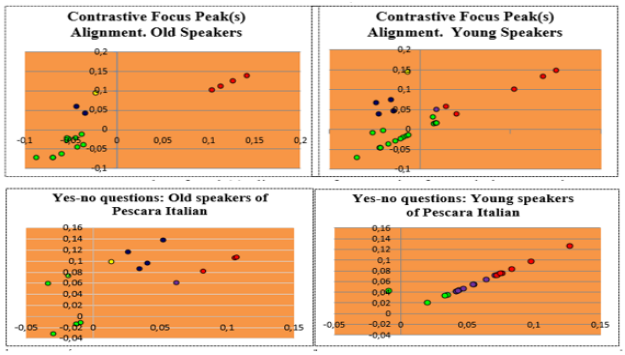


Figure 1. Pretonic peak (A, left), tonic peak (B, right), overlapping (C1, C2) and midway peak (D, bottom middle): highlighted squares bound the tonic syllable.

Table 1. Scatter plot: alignment of 1st peak (x axis) and 2nd peak (y axis), which are taken to coincide if there is only one peak (on the diagonal: types A, B, D).



	CF		YNQ	
	O	Y	O	Y
A	11	17	5	4
B	4	5	4	5
C1	1	4	1	0
C2	2	1	4	0
D	0	1	1	6

When coming to the youngsters, the evolution of this common distribution is sensitive to pragmatics: when associated to contrastive focus, patterns A) and

B) prevail, and intermediate patterns mildly increase, whereas in information-seeking questions type D) outnumbers A) and B) (6 occurrences, see Table 1)

Conclusion

As we deal with the same accents for the two sentence types, the asymmetry in change observed between the two sentence types can only depend on their pragmatics. As a possible reason, we posit that in questions speakers, needing information, are in a subordinate role and converge more (Gasoriek et al. 2015), leading to question contours phonetically converge into midway fudged forms, more than contrastive statements, which are a more assertive type. We suggest that researchers, whenever possible, carry on the documentation of all sentence types of a system, even if they want to analyze only one, in order to get more chances of coming across similar findings.

References

Alvord, S. 2010. Variation in Miami Cuban Spanish Interrogative Intonation. Hispánia 93 (2): 235-255.

Blum-Kulka, S., House, J., Kasper, G. 1989. Investigating cross-cultural pragmatics: an introductory overview. In S. Blum-Kulka, J. House, G. Kasper (Eds.), Cross-Cultural Pragmatics: Requests and Apologies. Norwood, NJ: Ablex, 1–34.

Chambers, J. K., Trudgill, P. 1980. Dialectology. Cambridge: CUP.

Gasoriek, J., Giles, H., Soliz, J. 2015. Accommodating new vistas. Langauge & Communication 41: 1-5.

Gili Fivela, B. et al. 2015. Intonational phonology of the regional varieties of Italian. In: S. Frota, P. Prieto (Eds.) Intonation in Romance. Oxford: OUP, 140-197.

Heffernan, K. 2006. Prosodic levelling during language shift: Okinawan approximations of Japanese pitch accent. Journal of Sociolinguistics, 10: 641-666.

O'Rourke, E. 2005. Intonation and language contact: a case study of two varieties of Peruvian Spanish. Ph.D thesis, University of Illinois at Urbana-Champaign.

Prieto, P., Borràs-Comes, J., Roseano, P. (Coords.) (2010-2014). Interactive Atlas of Romance Intonation. Web page: <http://prosodia.upf.edu/iari/>.

Romera, M., Elordieta, G. 2013. Prosodic accommodation in language contact: Spanish intonation in Majorca. International Journal of the Sociology of Language 221: 127-151.

Troncoso-Ruiz A., Elordieta, G. 2017. Prosodic accommodation and salience: The nuclear contours of Andalusian Spanish speakers in Asturias. Loquens 4 (2).

Vanrell, M., Feldausen, I, Astruc, Ll. The Discourse Completion Task in Romance prosody research: status quo and outlook. In: I. Feldhausen, J. Fließbach & M. Vanrell (Eds.) Methods in prosody: A Romance language perspective. Berlin: Language Science Press, 191-227.

Subject-object subextraction asymmetry in Russian

Daria Belova

Department of Theoretical and Applied Linguistics, Lomonosov Moscow State University, Russia
https://doi.org/10.36505/ExLing-2020/11/0012/000427

Abstract

Subject-object asymmetry in wh-subextraction is relevant in many languages, but the degree of subject opacity is not crosslinguistically uniform and can differ within one language. In (Polinsky et al. 2013) two factors are found statistically significant in Russian dependent clauses: the type of verbal structure (unaccusative / unergative / transitive verb) and the position of a subject relatively to a verb (preverbal / postverbal). In this paper we investigate whether the effect of these variables is preserved in Russian monopredicative independent clauses. Using experimental data, we show that different types of subjects do differ in their island properties, however, contra (Polinsky et al. 2013), in the preverbal position they are significantly more transparent to subextraction.

Keywords: subject-object asymmetry, wh-movement, word order, argument structure, topicality

Subject-object subextraction asymmetry

The opacity to subextraction of subject NPs relatively to object NPs was first captured by Ross (1967) and formulated as Sentential Subject Island Constraint (SSIC). Later this constraint was extended by Chomsky (1973) as Subject Condition and by Huang (1982) as Condition on Extraction Domain (CED): only complement position is transparent to subextraction.

However, the properties of subject islands are found not to be crosslinguistically uniform (see e.g. (Stepanov 2007). Subject NPs show different levels of opacity not only in different languages, but also within one language. These facts led to several theories explaining subject island constraints. Wexler and Culicover (1981) explain the asymmetry with the so-called *freezing effect*: a constituent becomes opaque after it has been moved out of its deriving position. Another approach is based on discourse-related *topic opacity*: only foci can undergo subextraction (see (Goldberg 2006)).

In this paper we explore two factors that are known to affect the subject NPs' opacity: verbal argument structure and word order. Subjects of different types of verbal structure are modelled as having different deriving position: a subject of unaccusative verbs is derived in a VP complement just like an object of a transitive verb, and subjects of transitive and unergative verbs are derived

in spec,vP. According to CED an unaccusative subject would be more transparent to subextraction than an unergative and a transitive subjects, but the mentioned approaches don't make such predictions. Change in word order is relevant because (i) non-basic word orders, if they are possible, demand movement from *in situ* position, what helps testing the *freezing* approach; (ii) the leftmost position is related to focus position, therefore, can provide evidence for the *topic opacity* approach.

Subject-object asymmetry in Russian

Previous research

Polinsky et al. (2013) were first to experimentally test the two factors mentioned above on Russian material. The authors investigated four types of NPs (subjects of unaccusative, unergative and transitive verbs and an object of a transitive verb) and two positions of NP relatively to a verb in Russian and found that (i) unaccusative subjects are indeed more transparent than subjects of the two other types, transitive subjects are the most opaque ones; (ii) objects in a preverbal position are opaque to subextraction. It is important to say that the researchers investigated only arguments within a dependent clause. However, Kush (2018) shows that the distance of movement is also relevant to island properties. For this reason, we have conducted our experimental investigation addressing the following research question: How the acceptability of subextraction from structurally different subjects and an object depends on the position of the subextraction site with respect to the verb in a monopredicative clause?

Current study

The experimental design included two independent variables: NP type (unaccusative, unergative, transitive subjects and a transitive object) and NP position (XV, VX where V stands for a verb and X stands for an NP). Half of the intransitive stimuli had a postverbal adjunct and half of the transitive stimuli had animate object. All the sentences began with wh-particle *čʹi* "whose", which was separated from a verb and its arguments by a time adverbial to make subextraction more obvious and avoid potential misinterpretations. One experimental block contained eight conditions. Three lexicalizations were made for each experimental condition, so there were 24 test stimuli in each experimental list and eight lists in total. Experimental lists also contained 24 fillers of two types: grammatical and ungrammatical. The ungrammatical ones included a filled gap in the place of wh-extraction. Respondents were asked to rate stimuli out of context on a Likert scale from 1 ("very bad sentence") to 7 ("very good sentence").

To check if non-basic word orders (VS for intransitive verbs, OSV/VOS/VSO for transitive verbs) could worsen the rating of the stimuli we have conducted

a pilot experiment. It had the same structure as the main experiment and shared the same lexicalizations but contained sentences without wh-extraction. The ungrammatical fillers contained mistakes in subject-verb agreement markers or incoherence between a time adjunct and verbal aspect.

103 Russian native speakers completed the pilot experiment (53 females), ranging from age 13 to 67 (mean = 33). 171 respondents participated in the main experiment (86 females), ranging from age 16 to 64 (mean = 34). The ratings of each respondent in both experiments were normalized to minimize personal tendencies in using the scale.

The results of the pilot experiment have shown that non-basic word orders do worsen ratings, so it provides us a baseline to estimate the role of these factors un the main experiment.

Finally, the main experiment showed the following results. According to ANOVA both NP type and NP position have significant effect on ratings. Besides, (i) the preverbal position is significantly less opaque to subextraction than the postverbal position for each type of NP, including the object; (ii) in the preverbal position all subjects are rated the same and the object is rated lower; (iii) in the postverbal position unaccusative subject is the most(?) transparent, unergative subject is rated lower and both arguments of transitive verb are rated almost equally as the less transparent (the statistical difference between them is marginal):

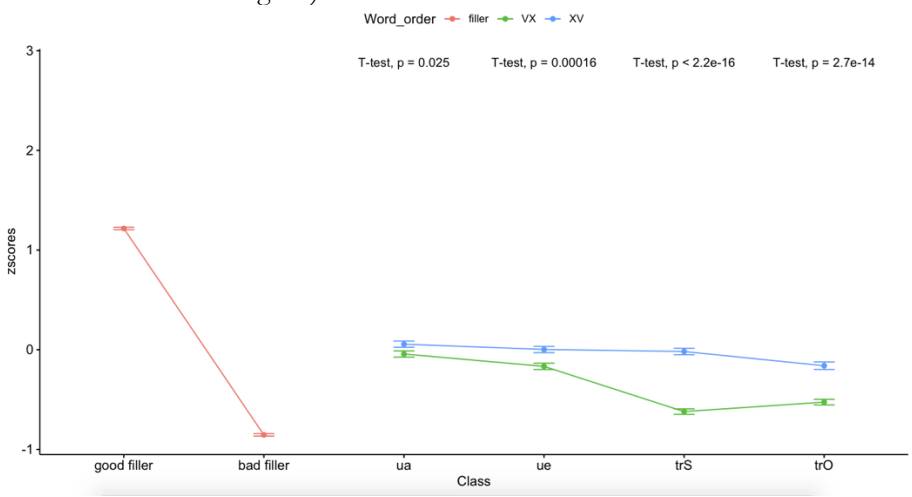


Figure 1. The interaction plot with the factors *NP type* and NP *position* and the ratings of the two types of fillers. On the top of the plot the results of the Student's t-test application to pairwise comparisons of two positions for each verb type are presented. *Ua*, *ue*, *trS*, *trO* stand for *unaccusative*, *unergative* and *transitive* subjects and object.

These results are remarkable from two points of view. Firstly, our findings contrast with the results of (Polinsky et al. 2013): in our experiment the object

NP is far less opaque and the position of NP has much stronger effect on the ratings of subextraction from every type of subject. Secondly, the obtained results are not consistent with predictions of the three presented approaches. Nevertheless, our results support the hierarchy found in (Polinsky et al. 2013): the unaccusative subextraction gets the highest ratings and subextraction from the transitive subjects gets the lowest ones. Thus, our results suggest that the distance of movement is a relevant factor for Russian subject island properties.

References

Chomsky, N. 1973. Conditions on transformations. In Anderson, S., Kiparsky, P. (eds). 1973. A festschrift for Morris Halle, 232-286. New York: Holt: Rinehart & Winston.

Goldberg, A. 2006. Constructions at Work: The nature of generalization in language. Oxford: Oxford University Press.

Huang, C.-T.J. 1982. Logical relations in Chinese and the theory of grammar. Ph.D. dis. Massachusetts University of Technology.

Kush, D., Lohndal, T., Sprouse, J. 2018. Investigating variation in island effects. A case study of Norwegian wh-extraction. Natural language and linguistic theory 36. 743-779.

Polinsky, M. et al. Subject islands are different. In Sprouse, J., Hornstein, N. (eds). 2013. Experimental syntax and island effects, 286-309. Cambridge: Cambridge University Press.

Ross, J.R. 1967. Constraints on variables in syntax. Ph.D. dis. Massachusetts Institute of Technology.

Stepanov, A. 2007. The end of CED? Minimalism and extraction domain. Syntax 10(1), 80-126.

Wexler, K., Culicover, P. 1981. Formal principles of language acquisition. Cambridge, Mass.: MIT Press.

Well-established monolingual literacy predictors in bilinguals

Insiya Bhalloo[1,2], Kai Ian Leung[1,2], Monika Molnar[1,2]
[1]Bilingual and Multilingual Development Lab, University of Toronto, Canada
[2]Rehabilitation Sciences Institute, University of Toronto, Canada
https://doi.org/10.36505/ExLing-2020/11/0013/000428

Abstract

An important component of early reading intervention is effective literacy screening tools. Literacy precursor screening tools have been primarily developed for early identification and remediation of potential reading difficulties in monolingual English-speaking children, despite the significant proportion of bilingual children worldwide. This systematic literature review examines whether the precursor literacy skills commonly used in monolingual English-speaking children have been assessed and found to predict later reading skills in simultaneous bilingual children. Our findings demonstrate that the nine major literacy precursors identified in monolingual children also significantly correlate with reading performance in simultaneous bilingual children. These nine literacy precursors are phonological awareness, letter knowledge, serial recall, oral language comprehension, vocabulary, grammar, memory, non-verbal intelligence and word decoding.

Keywords: child literacy, education, literacy precursors, bilingualism, multilingualism

Introduction

Globally, reading difficulties are common in young children, with 56% of primary school-aged children unable to meet age-appropriate reading proficiency levels (UNESCO, 2017). Literacy precursors are language and cognitive skills, developed prior to a child receiving formal literacy instruction, that predict child success in future reading acquisition (Hjetland, Brinchmann, Scherer, & Melby-Lervåg, 2017). Identifying literacy precursors is crucial for early intervention of children at risk for future word decoding and reading comprehension difficulties, as these children typically exhibit deficits in precursor literacy skills. Word decoding is the ability to first identify individual phonemes present in a written word and then sequence combinations of phoneme (sound) - grapheme (symbol) associations to pronounce/read a printed word, while reading comprehension is the ability to recognize and then understand written text. Based on previous systematic reviews conducted on monolingual literacy development (i.e. Hjetland et al., 2017), the following nine critical precursors of future word decoding and reading comprehension have been identified in monolingual children: (i) phonological awareness, (ii) letter knowledge, (iii) serial recall, (iv) oral language comprehension, (v) receptive and

expressive vocabulary, (vi) grammar (including morphological and syntactic awareness), (vii) memory, (viii) non-verbal intelligence, and (ix) word decoding.

The majority of child literacy development research has focused on monolingual children; however, approximately 50% of children worldwide grow up in a bilingual learning environment (e.g., Ryan, 2013). It is important to examine whether the literacy precursors, identified in monolingual children, have been assessed and found to predict reading development in bilingual children. As evident by previous research, differences have been observed in cognitive, speech, language and literacy development between monolingual and bilingual children (Thomas-Sunesson, Hakuta, & Bialystok, 2018). Such a knowledge synthesis that focuses on bilingual children is therefore necessary in order to support reading abilities in both bilingual languages, due to these observed oral language and cognitive-linguistic differences.

Purpose of the study

This systematic review investigates whether the nine literacy precursor skills (as established in monolingual children) have been assessed and found to predict later reading performance, as measured by word decoding and reading comprehension, in one or both languages of simultaneous bilingual children.

Methods

We conducted a database search of the nine literacy precursor skills in relation to the two literacy outcome measures, word decoding and reading comprehension. We searched four linguistics and education-based electronic databases: Linguistics and Language Behavior Abstracts (LLBA), Educational Resources Information Center (ERIC), MLA International Bibliography and PsycINFO (ProQuest). The database search was conducted based on the PICO framework: Population (of typically-developing simultaneous bilingual children aged 12 years or younger), Intervention (the type of literacy precursor assessed in relation to reading), Comparison (not applicable, as we did not specify a mandatory monolingual, sequential bilingual or second language learner age-matched comparison group) and Outcome (the type of reading-based literacy outcome measure assessed).

In our review, we only included peer-reviewed journal publications that assessed precursor literacy skills in relation to a reading-based literacy outcome measure in typically-developing simultaneous bilinguals (i.e. they were exposed to both languages before the age of 3) who were 12 years of age or younger. To account for the risk of bias, we excluded studies that did not meet the Critical Appraisal Checklist for Quasi-Experimental Studies (Joanna Briggs Institute, 2017). Additionally, we excluded studies that did not indicate sufficient child language background and parental socio-economic status information.

Results and discussion

The database search resulted in a total of 2433 studies, with 1464 duplicate items. After screening for duplicates, 969 abstracts were identified as relevant for title and abstract-level screening. Two reviewers (IB and KIL) independently screened and identified 248 abstracts that met the inclusion criteria. After conducting the full-text screening and critical appraisal of these studies, 18 studies met the inclusion criteria. In addition, we manually searched the reference list of these 18 papers as well as the previously excluded non-primary studies. Using Google Scholar, we also screened papers that cited these studies. This resulted in 22 additional studies (after the initial title and abstract-level screening). After conducting full-text screening of these additional studies, 11 additional studies met the inclusion and critical appraisal criteria. In total, 29 studies (18 from the original full-text screening and 11 from the additional screening) met the inclusion and critical appraisal criteria and were included in this systematic literature review.

A total of 23/29 reviewed bilingual studies assessed these nine monolingual precursors in simultaneous bilingual children: phonological awareness (13 studies), letter knowledge (2 studies), serial recall (3 studies), oral language comprehension (3 studies), vocabulary (14 studies), grammar (6 studies), memory (2 studies), non-verbal intelligence (1 study), and word decoding (4 studies). The most commonly assessed literacy precursors in the reviewed bilingual studies were vocabulary, phonological awareness, grammar and word decoding. The literacy precursors with the greatest number of significant, as compared to non-significant, correlations were vocabulary, phonological awareness, oral language comprehension, letter knowledge and non-verbal intelligence (see Table 1).

The significant correlations between these nine literacy precursors and the literacy outcome measures indicate that similar literacy precursor screening tools and reading outcome measures, developed for monolingual children, may also be used for assessing simultaneous bilingual children. However, it is important that speech-language pathologists and educators do not generalize child literacy research findings on monolingual English-speaking children to monolingual and bilingual children from diverse language backgrounds. This review demonstrates that the specific type of literacy precursor assessed impacts the type and strength of within- and across-language correlations between given literacy precursors and literacy outcome measures. Further research on developing language-specific literacy precursor and literacy outcome measures, based on the language(s) of assessment's linguistic properties is needed to effectively assess child reading abilities and provide early reading intervention for monolingual and bilingual populations in all their spoken languages. Linguistic diversity in child literacy research and developing appropriate language-specific screening tools for various linguistic populations

will enable early identification and intervention of potential reading difficulties in both heritage and societal languages.

Table 1. Significant and non-significant within- and across-language correlations between literacy precursors (indicated on the left) and literacy outcome measures in monolingual (Hjetland et al., 2017) and simultaneous bilingual children. *Note.* M= monolinguals; SB= simultaneous bilinguals; sig.= significant; N.S.= not-significant; WL= within-language; AL=across-language.

Literacy Precursor (n= number of studies)	M	SB			
	Sig.	Sig. WL	Sig. AL	Sig. WL & AL	N.S
Phonological Awareness (n=13)	Y	n=10		n=3	
Letter Knowledge (n=2)	Y	n=1		n=1	
Serial Recall (n=3)	Y	n=1		n=1	n=1
Oral Language Comprehension (n=3)	Y	n=1		n=2	
Vocabulary (n=14)	Y	n=8	n=1	n=5	
Grammar (n=6)	Y	n=3		n=2	n=1
Memory (n=2)	Y		n=1		n=1
Non-Verbal Intelligence (n=1)	Y			n=1	
Word Decoding (n=4)	Y	n=3			n=1

References

Hjetland, H. N., Brinchmann, E. I., Scherer, R., Melby-Lervåg, M. 2017. Preschool predictors of later reading comprehension ability: A systematic review. Campbell Systematic Reviews, 14, 1-156.

Joanna Briggs Institute. 2017. Critical appraisal checklist for quasi-experimental studies (non-randomized experimental studies).

Ryan, C. 2013. Language use in the United States: 2011. American Community Survey Reports, ACS-22. U.S. Census Bureau, Washington, DC.

Thomas-Sunesson, D., Hakuta, K., Bialystok, E. 2018. Degree of bilingualism modifies executive control in Hispanic children in the USA. International journal of Bilingual Education and Bilingualism, 21(2), 197-206.

United Nations Educational, Scientific and Cultural Organization (UNESCO). 2017. More than one-half of children and adolescents are not learning worldwide.

Duration measurement of vocants in infants with Cochlear Implants

Paris Binos[1], Chryssoula Thodi[2], George Psillas[3], Paris Vogazianos[2]
[1]Department of Rehabilitation Sciences, Cyprus University of Technology, Cyprus
[2]Department of Health Sciences, European University Cyprus, Cyprus
[3]Department of Medicine, Aristotle University of Thessaloniki, Greece
https://doi.org/10.36505/ExLing-2020/11/0014/000429

Abstract

Vocants are precursors to speech and are facially neutral. The presence of these speech-like vocalizations was evident during the precursors to mature phonology called "protophones". The prosodic feature of duration of the nuclei plays a crucial role in the shift of prelexical to mature speech, since speech intelligibility is closely related to the control of duration. The aim of this work is to determine whether cochlear implants (CIs) positively trigger language acquisition and the development of verbal skills. Recent literature findings are compared and discussed with the performance of two Greek congenitally hearing-impaired infants who were matched with three normal-hearing (NH) infants. This work highlighted an important weakness of the prosodic abilities of young infants with CIs.

Keywords: vocants, prosody, duration, deaf, cochlear implant

Introduction

Language and communication are primarily based on hearing. Hearing deprivation interferes with acquisition of crucial acoustic cues, limits word comprehension, and lowers oral language performance it . Although paediatric cochlear implantation provides measurable benefits on language perception, persistent high heterogeneity of findings prevents a thorough understanding of the process (Spencer, Marschark, Spencer, 2011).

Vowel duration during babbling is crucially important for future speech and language development: shorter vocal duration is related to mature adult-like vocalizations (Kent, Forner, 1980). There is an extensive work related to mid-front, low-front and central vowels of children with normal hearing (NH) (Fernald 2004, Katz, Cohn, Moore, 1996, Kitamura, Lam, 2009) during the babbling stage but there are only few reported studies focused on vocant prosodic features (quasivowels and full vowels) of infants with severe-to-profound hearing loss despite their clinical importance (Binos, Thodi, Vogazianos, Psillas, Constantinidis, 2020).

Aim of the study

The aim of this literature overview was to present recent findings on suprasegmental duration features in young CI recipients based on very early speech transitions called vocants, and to compare a recent study based on Greek language (Binos et al. 2020) other studies exploring similar phenomena.

Method

We reviewed published peer-reviewed studies from the last ten years identified via a PubMed/Medline search. The keyword combination "vowel duration" revealed seventeen papers and "vowel duration cochlear implants" revealed two papers. . Inclusion criteria included whether the papers were research reports, clinical trials, meta-analysis, randomized controlled trials, reviews and systematic reviews and only if a full-text access was possible; the papersshould report on infants with CI with speech analysis in the babbling stage. This focus on studies from the last ten years was dictated by recent technological developments that shaped cochlear implants capabilities.

Based on our inclusion criteria studies related to adults or normal-hearing participants were excluded. The papers included (Neumeyer, Harrington, Draxler, 2010, VanDam, Ide-Helvie, Moeller, 2011).

The Greek language study (Binos et al, 2020). Binos et al., (2020), analyzed vocant productions based on spontaneous vocalizations using a combination of acoustical and auditory characteristics. Reflexive or vegetative sounds were excluded.Our CI participants had deafness of unknown etiology and no other disabilities and were enrolled in auditory-oral programs.

Prior to implantation, the infants of the Greek study had an average unaided hearing loss of 95–110 dBHL, while the post implantation pure-tone audiometry (PTA) was 10–38 dBHL. A wideband spectrographic analysis was performed selecting segments from voice onset to the loss of audible signal and the decrease of higher formant energy. To match children for hearing experience, infants with CIs were chronologically older and more developmentally mature than NH peers. Due to the high F_0 of infants' speech, a large analysis bandwidth (450 Hz) was used for the spectrographic analysis. The vocant onset and offset were measured via wideband spectrographic analysis from the voicing onset (first glottal pulse) to the loss of audible signal and the decrease in the energy of higher formants (F2). The decrease of energy was visualized through wideband spectrograms where a lowering of F2 was visible while the oral cavity was enlarged because of tongue retraction (Binos et al, 2020).

Results

During the Binos et al (2020) analysis the Kolmogorov-Smirnov test was used for sample sizes greater than 50 and the Shapiro-Wilk test for samples smaller

than 50 to check for data distribution. The data analyzed did not follow the normal distribution for the duration variable. The Levene Test of Equality of Variance was carried out for the duration variable of the groups and the results showed no homogeneity of variances for the duration variable making the Median test more appropriate than the Mann Whitney test. The results (δ = 512.5, χ^2 = 29.020, $p < .001$) displayed a major weakness for the CI participants on the duration performance. The average vocant duration for NH infants was significantly lower from the CI infants.

Discussion

The studies of VanDam, Ide-Helvie and Moeller (2011) and Yang, Brown, Fox et al (2015) agree with the Greek study since the lower the chronological age the longer vocal duration. The VanDam et al (2011) study included groups of CI users ages8-9 years old and the duration of vowels was found longer but only for typical development children when the vowels were followed by consonants (type CVC). In the Yang et al (2015) study pre-lingually deafened CI children (2.9-8.3 years old) produced seven Mandarin vowels longer than the NH children did. Longer duration measured for the isolated vowels /e, α/, /i α/ of CI users in German (Neumeyer et al, 2010). The Greek study broadens these findings to younger CI individuals.

Longer vocant duration in young CI participants is considered a poor predictor for speech production skill development. Recent studies associate shorter duration of vowels with greater speech maturity and greater intelligibility. The clinical implications of our finding are critical for articulatory positions and acoustic outputs, as habilitation models for young CIs should be adjusted accordingly.

Acknowledgements

The authors warmly thank all families that participated in the Greek study. We also thank the University Hospital of AHEPA in Thessaloniki, Greece. Our special thanks go to the ENT Clinic for their continuous support and participation.

References

Binos, P., Thodi, C., Vogazianos, P., Psillas, G., Constantinidis, J. 2020. An acoustic and auditory analysis of vocants in infants with cochlear implants. Logopedics Phoniatrics Vocology, DOI: 10.1080/14015439.2020.1724325.

Fernald, A. 2004. Hearing, listening, and understanding auditory development in infancy. In *Blackwell handbook of infant development*. London: Blackwell Publishing: 35–70.

Neumeyer, V., Harrington, J., Draxler, C. 2010. An acoustic analysis ofthe vowel space in young and old cochlear-implant speakers. Clin Linguist Phon 24(9), 734–741.

Katz, G., Cohn, J., Moore, C. 1996. A combination of vocal f0 dynamic and summary features discriminates be- tween three pragmatic categories of infant-directed speech. Child Dev 67, 205–217.

Kent, R., Forner, L. 1980. Speech segment durations in sentence recitations by children and adults. J Phon 8(2), 157–168.

Kitamura, C., Lam, C. 2009. Age-specific preferences for infant-directed affective intent. Infancy 14(1), 77–100.

Spencer, P.E., Marschark, M., Spencer, L.J. 2011. Cochlear implants: Advances, issues and implications. In *The Oxford Handbook of Deaf Studies, Language and Education*. (eds.) (vol. 1, 2nd edn). New York, NY: Oxford University Press: 452-471.

Yang, J., Brown, E., Fox, R., et al. 2015. Acoustic properties of vowel production in prelingually deafened Mandarin-speaking children with cochlear implants. J Acoust Soc Am. 138(5), 2791–2799.

The outcomes of neglecting native language teaching

Natia Botkoveli
Vytautas Magnus University, Lithuania
https://doi.org/10.36505/ExLing-2020/11/0016/000431

Background information

The republic of Georgia is one of the most linguistically diverse regions in the Caucasus. As other multilingual countries Georgia is also facing difficult decisions concerning which national or official languages to use in administrative and education sectors. On the surface, the ethnic situation in Georgia seems to be well managed; however it cannot be said about the language situation. One of the issues that leads to confusion and is possibly the most discussed issue by Georgian linguists is the undefined status of the West-Kartvelian languages, i.e. Megrelian and Svan. Svans, with a population of about 5,500, are mostly located in the high valleys along the Enguri and Tskhenistskali rivers in Western Georgia. Megrelians, around half a million, live in the capital and in the western Georgian lowlands (Jorbenadze 1991:15).

Investigating the status of Megrelian and Svan is a sensitive matter. The most discussed question in contemporary Georgian sociolinguistics is the qualification of these linguistic codes as languages or dialects. Some scientists (Gvantseladze 2006; Putkaradze 2010) consider them to be dialects of Georgian. Their opinion is mainly based on two arguments (Putkaradze 2010: 138, 139), however Wheatly (2009) argues that using purely objective criteria it is hard to argue that Megrelian and Svan are dialects of Georgian since they are not mutually comprehensible with Georgian (Wheatly 2009: 28).

Megrelian and Svan never gained the status of the permanent literary language during the Soviet period, a number of books, journals and newspapers were produced (mostly in Megrelian) (Jorbenadze1991: 22), which has not happened in cases of other dialects of Georgian (Dragojevic et al. 2014: 20). According to the leading opinion in Kartvelology, both Megrelian and Svan are languages, related to Georgian and these three languages comprise the same language family.

Methodology

The practical aim of this paper is to identify the existing challenges of the current education and language management and demonstrate the factors that are caused by the omission Megrelian and Svan in the education policy. This study is a first attempt in the Georgian linguistic tradition to address the current

language policy and language management from the different points of views in order to investigate:

(1) How does the current language management affect the unrecognized West-Kartvelian languages?
(2) What are the main obstacles for Georgian policymakers to achieve a multilingual policy?

The main objective of the study is to identify the difficulties of the current language management problems and support the development of Georgian language management, which will represent and sustain the diversity of the languages spoken in the country.

A qualitative semi-structured interview method was employed to elucidate the research questions and collect the data from the respondents involved in the current language management process. The obtained data represents the ideology of academia, government and language activists. I believe that understanding and analyzing different and often times contrasting views will contribute to a better understanding of the issues and finding solutions.

The study was conducted in Tbilisi, the capital of Georgia, which gathers professionals from around the country, including Samegrelo and Svaneti. All interviews were carried out in an indoor location, typically in a house or in the office. In each case, there was one interviewer and one respondent only. The interviews were carried out as informal conversations to ensure a relaxed atmosphere so that the interviewees felt comfortable to talk freely. In order to ensure the maximum of their openness and privacy, they were informed in advance that none of their personal information was going to be heard in the audio or shown in the transcript.

The criteria for choosing the interviewees was based on the initial aims of the study to approaches the topic from three different viewpoints:

(i) Authorities, which made it possible to identify the existing problems of language management – 4 interviews totalling 128.72 minutes;
(ii) Researchers of Svan/Megrelian languages, who discussed the challenges academia members face due to lack of interest towards the West-Kartvelian languages – 4 interviews totalling 136.48 minutes.
(iii) Language activists, who experience the shortcomings of the current policy– 3 interviews totalling 184.29 minutes.

After transcribing the interviews verbatim, data were analyzed using content analysis. While the content analysis can be carried out at different levels of depth, my interest was an in-depth and comprehensive understanding of the problems caused by the omission of Megrelian and Svan in the language management processes.

Results and discussion

The study demonstrated that current education policy supposedly leads to the endangerment of the West-Kartvelian languages. Respondents claim, that the

modern linguistic management can be a reason for problem, which was confirmed on the international level as well.

According to the document titled "Language Vitality and Endangerment" (www.unesco.org) compiled by UNESCO in 2003, Megrelian and Svan languages are described as Definitely Endangered. According to Atlas of the World's Languages in Danger, children no longer learn the languages as mother tongue in the home.

Noticing the problem of endangerment was a turning point for the all language activists I interviewed within this research. Respondent #6, shared his observations that even his parents are unable to speak the "same" Megrelian as his grandmother used to speak.

Respondent #3, who has rich experience of fieldwork in Samegrelo, is convinced that for all children born in Samegrelo in the recent years, the L1 is Georgian. She shared that some Megrelian parents are even proudly claiming that their children barely speak Megrelian and the new generations are already fluently expressing themselves in Georgian. Very similar information was provided by the respondent #5, who shared her experience with Svan children in the high mountainous villages of Svaneti:

> "At the very beginning I was thinking that it was only the children from the bigger settlement who did not speak proper Svan, but after observing the villagers, I noticed that they too were using Georgian only… I was shocked!"

I asked my respondents to identify the reasons for this tendency. Respondent #3 believes that there are two main factors that cause this problem: (i) the lack of prestige of the West-Kartvelian languages; (ii) the negative attitudes towards the accent and intonation Megrelian and Svan speakers have while speaking in Georgian.

During this research the grass-root officials were asked to discuss the reasons hindering the introduction of Megrelian and Svan at the school level. It turned out that Georgia lacks a language policy document. Respondent #9 explained that the State Language Department is working on the new language policy document, which will include a ten year strategy for language management.

Meanwhile, young Megrelian and Svan speakers are discouraged from studying and researching the West-Karvelian languages. Researcher #5 suggests that the factors, which lead to a lack of motivation to learn or deepen their knowledge in Megrelian and Svan are easy to explain as there are no career opportunities in these languages:

> "Imagine someone born in the capital, why would they "waste" their time on learning Megrelian and Svan when they can learn something more useful like English or German?"

I was given a chance to hear from the those who do not admit that the problem of endangerment is real. Respondent #9, who occupies one of the head positions in the State Language Department, argues that the criteria for determining the endangered languages lack accuracy. She claims that several years ago Megrelian and Svan were not included in the UNESCO's list of endangered languages. According to her there was no clear evidence that the number of the West-Kartvelian language speakers decreased drastically in the last two decades. She suggests that some criteria for assessing the level of endangerment are somehow misleading:

> "If we compare Megrelian and Svan we will notice that they have quite different qualitative and quantitative indices of their native speakers. Megrelian is being spoken by more than a million speakers, whereas Svan is being used by less than 3000 people. It is obvious which one is endangered, but both are in the list."

The discussion regarding the endangerment of the West-Kartvelian Languages highlights the large diversity in attitudes and beliefs. The key take-away is that the respondents had strong feelings related to the endangerment of Megrelian and Svan languages. One of the main reasons for this topic being so touchy might be the fact that the majority (7) of my respondents have Megrelian and Svan backgrounds. The officials have expressed controversial ideas on this issue, which once again demonstrated the need for discussions where all parties could be involved.

In summary, for the majority of the participants limited education in Megrelian and Svan seems to be a serious problem as well as a possible reason for the language endangerment. Despite the language activists' strong desire to introduce the West-Kartvelian languages in schools some researchers and authorities do not approve of this idea. Supposedly, until adopting the new language and education policies the future of education accessible in Megrelian and Svan remains undecided.

References

Dragojevic, Marko and Berglund, Christofer and Blauvelt, Timothy. 2014. Attitudes Towards Tbilisi and Mingrelian-Accented Georgian Among Georgian Youth: On the Road to Linguistic Homogenization? Tbilisi: Iliauni Press.

Jorbenadze, Besarion. 1991. The Kartvelian Languages and Dialects. Tbilisi: Mecniereba.

Putkaradze, Tariel and Dadiani, Eka and Sherozia, Revaz. 2010. European Charter for Regional or Minority Languages and Georgia (Sakartvelo). Kutaisi: Kutaisi Publishing Center.

Wheatley, Jonathan. 2009. Georgia and the European Charter for Regional or Minority Languages Flensburg: European Centre for Minority Issues (ECMI).

Atlas of the World's Languages in Danger 2011.
http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/pdf/aboutEndangeredLanguages-WV-EN-1.pdf (date of retrieval: 20 August 2020).

Tone-consonant co-occurrence probability in Mandarin

Chiung-Yu Chang, Feng-fan Hsieh
Institute of Linguistics, National Tsing Hua University, Taiwan
https://doi.org/10.36505/ExLing-2020/11/0017/000432

Abstract

This study investigates how statistical regularity concerning the distribution of lexical tones and consonantal onsets in Mandarin, formulated as the "tone-consonant co-occurrence probability (TCCP)" here, influences results of wordlikeness judgment task. Native speakers were asked to rate the wordlikeness of monosyllabic real words and pseudowords with existing segmental combinations. Overall, real words with high probability were considered more wordlike than those with low probability. On the other hand, the probability effect was not significant on the well-formedness ratings of the pseudowords. These findings suggest that speakers are sensitive to the tone-consonant co-occurrence patterns, which follow gradual tendencies rather than an "all-or-nothing" manner, but such sensitivity is probably limited to existing forms and cannot be extended to hypothetical ones.

Keywords: phonotactic probability, wordlikeness, consonant-tone interaction, pseudowords, Mandarin

Introduction

It has been long noted that some consonantal onsets rarely co-occur with certain lexical tones in Mandarin Chinese. As a result of the historical tone split, unaspirated obstruent onsets seldom appear on a Tone 2 (mid-rising) syllable, and sonorant onsets rarely occur on a Tone 1 (high-level) syllable (Tung 1970). Nevertheless, there is still a fair amount of counterexamples to these robust tendencies, such as [pa2] 'to pull' and [ma1] 'mother'. Therefore, it is safe to say that the above co-occurrence restrictions are not categorical in nature.

To better capture the probabilistic nature of the tone-consonant co-occurrence restrictions, we propose a new metric that is analogous to conditional phonotactic probability. For a given consonantal onset X and a tone Y, tone-consonant co-occurrence probability (TCCP) for the combination of X∩Y is operationally defined as the type frequency of the co-occurrence of X∩Y divided by the type frequency of X. Further details of calculating TCCP will be provided in the method section.

In addition, we conducted a wordlikeness judgment experiment to test whether native speakers possess the knowledge of Mandarin tone-consonant association. In English, sequence probabilities of segments have been shown to exert unique effects on the wordlikeness of pseudowords after partialling out

neighborhood effects (Bailey and Hahn 2001). By the same token, this study sets out to explore whether TCCP, which involves both segmental and suprasegmental entities, influences wordlikeness in a similar fashion.

Method

Participants

17 native speaker of Taiwanese Mandarin (age range: 20-28, 7 female, without speech or hearing impairment) participated in the experiment. All of them signed an informed consent and were paid for their participation.

Materials

Type frequency was approximated as the number of lexical entries that contain the particular sound or sound combination in the Concise Mandarin Chinese Dictionary (National Languages Committee 2000). The phoneme inventory used for calculation was the same as the phonological representation system adopted by this dictionary, i.e., Mandarin Phonetic Symbols. The resulting TCCPs ranged from 0.002 (for [ʐ]∩Tone 1) to 0.516 (for [k]∩Tone 1). The lowest 14 and highest 15 onset-tone combinations were considered as having low and high TCCPs, respectively.

42 pseudowords with existing segmental combinations (e.g., [tau2]) and 42 real monosyllabic words are chosen as test materials. Half of the pseudowords and real words have a high TCCP, while the other half have a low TCCP. Each item was divided into two or three parts according to the schema — consonantal onset, nucleus, nasal coda (if available), and tone. Neighborhood density (ND) was measured as the sum of the number of monosyllables that have a Levenshtein edit distance of 1 from the stimuli and were entered as a covariate in the statistical analysis.

Procedure

Prompted with visual stimuli written in Mandarin Phonetic Symbols, participants were asked to try producing the (pseudo)words and evaluate their wordlikeness based on a 7-point Likert scale, with 7 being "the most wordlike".

Results

As illustrated in Figure 1, real words with a low TCCP received considerably more low scores than those with a high TCCP. The distinction between high and low TCCPs was not obvious for pseudowords.

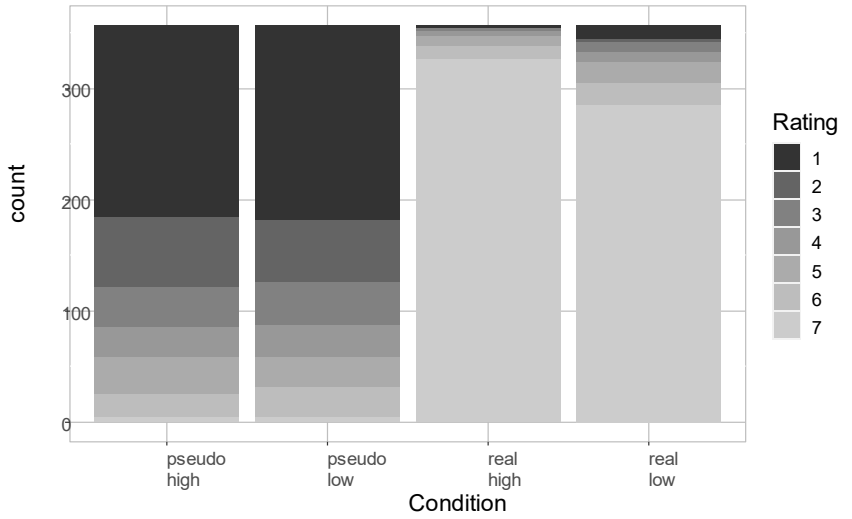


Figure 2. Rating results, where "pseudo" stands for "pseudowords".

The data was analyzed with a series of cumulative link mixed models built by the ordinal package (Christensen 2018) in the R software (R Core Team 2017). All models contained random intercepts for Participants and Items. Lexicality, TCCP, ND, and the interactions among these three factors were entered as fixed effects. Starting from the highest-order term, likelihood ratio tests were performed to assess the contribution of each predictor. The most parsimonious model contained TCCP and its two-way interactions with ND and Lexicality. The significance of each term was confirmed by likelihood ratio tests (all p-values < 0.05).

Separate models were then constructed for the data of real words and pseudowords. Low TCCP was coded as the baseline. Using the method of model selection described above, the best model for real words was the one with only the TCCP effect (β = -1.30, SE = 0.53, Z = -2.44, p = 0.015), and that for pseudowords incorporated only the ND effect (β = 0.07, SE = 0.02, Z = 1.91, p = 0.056).

Discussion and conclusion

In this study, we presented a lexical metric to quantify the affinity of specific types of tones and consonants, which helped explain the results of wordlikeness judgment task to a larger extent in the statistic models. The co-occurrence probability is thus a useful method for describing the relationship between consonants and tones.

The significant effect of TCCP on wordlikeness indicates that native speakers are sensitive to the gradient nature of tone-consonant co-occurrence in the lexicon. Interestingly, TCCP interacts with Lexicality and ND. Yao and

Sharma (2017) has reported the influence of ND on the reaction time of real words and pseudowords in a lexical decision task. The fitted regression coefficient of the ND effect observed was generally larger for pseudowords than real words in their experimental results. This is consistent with the current findings that the ND effect was present in the model of pseudowords but not real words. On the other hand, the effect of TCCP was confined to real words, unlike the results in English (Bailey and Hahn 2001). The limited role of TCCP might be attributed to the less significant role of lexical tones in comparison with segments in Mandarin, as evidenced by the lower fidelity of suprasegmental properties in loanword adaptation, as well as many psycholinguistic studies (e.g., Ye and Connine 1999).

A limitation of this study is that only a relatively small set of phoneme combinations are included in the materials. Data collected for more items and from more participants would help us establish a more comprehensive picture of the determinant factors of wordlikeness, including TCCP and other lexical variables.

Acknowledgements

This work was supported by the grant MOST106-2420-H007-010 from Ministry of Science and Technology, Taiwan.

References

Bailey, T.M., Hahn, U. 2001. Determinants of wordlikeness: Phonotactics or lexical neighborhoods? Journal of Memory and Language 44, 568-591.

Christensen, R.H.B. 2018. Ordinal - Regression Models for Ordinal Data. R package version 2018.4-19. http://www.cran.r-project.org/package=ordinal/.

National Languages Committee (ed.) 2000. Concise Mandarin Chinese Dictionary. Taipei, Ministry of Education.

R Core Team, 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Tung, T'.-H. 1970. Hanyu Yinyunxue [Chinese Phonology]. Taipei, Student Book.

Yao, Y., Sharma, B. 2017. What is in the neighborhood of a tonal syllable? Evidence from auditory lexical decision in Mandarin Chinese. In Farrell, P. (eds.) 2017, Proceedings of the Linguistic Society of America 2, 45:1-14, https://doi.org/10.3765/plsa.v2i0.4090

Ye, Y., Connine, C. M. 1999. Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes 14*(5-6), 609-630.

The phonation types in Fuzhou Chinese

Changhe Chen
Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, China
https://doi.org/10.36505/ExLing-2020/11/0018/000433

Abstract

This pilot study provides an acoustic description of the phonation types in Fuzhou Chinese. Speech samples from 5 native speakers show that tones [21 242 24] are breathy, while [44 32 4] are mainly modal and [51] is modal-breathy. Acoustic measure HNR35 can distinguish these phonation types, while H1*-A1* can only differentiate [24] and the end of [51] from other tones.

Keywords: Fuzhou Chinese, phonation type, voice quality, tone-vowel interaction

Introduction

Fuzhou Chinese is a variety of Eastern Min spoken in southeastern China. It has seven citation tones and is known for its tone-vowel interaction. For instance, [i] (only carrying citation tones [44 51 32 4][2]) - [ei]/[ɛi] ([21 242 24]), [y] - [øy]/[ɵy], [e] - [a], [ø] - [ɔ], [o] - [ɔ], [u] - [ɔu], [øy] - [ɔy] and so on. In tone sandhi position, the lower or diphthongal vowels will become the higher or monophthongal ones. Donohue (2017) reports that tones [21 242 24] are creaky and explains such interaction from the perspective of phonetic enhancements. However, Esposito (2006) analyses recordings of a male speaker (aged 27 in 2004) and shows that tone [21] is produced with breathiness which can be distinguished from modal voice by measures CPP, H1-H2, H1-A1, H1-A2 and so on. In view of scanty treatises and inconsistent findings, this pilot study aims at providing an acoustic description of seven citations tones in terms of their phonation types.

Methods

Speech samples from 5 native speakers (3 females) aged above 60 were recorded[3]. They read a word list for five times at the normal speech rate. From that list, the vocalic parts in [tsʰø^{44}], [e^{51}], [tøy^{32}], [ei^{21}], [tøy^{242}], [ɔʔ^{24}], [oʔ^{4}][4] are analysed regarding values of HNR35 (Harmonic-to-Noise Ratio 0-3500 Hz) and H1*-A1* (the amplitude difference between the first harmonic and the harmonic nearest to the first formant) obtained by using VoiceSauce (Shue et al., 2011). The second measure is corrected for formant frequencies and bandwidths in VoiceSauce, so data from different vowels can be compared. Ten time points are evenly sampled throughout the whole duration excluding the onset and end[5].

Results and Conclusions

HNR measures noise in signal, and both creakiness and breathiness can lead to lower values (Garellek, 2012). Thus, prior to quantitative analysis, the spectrograms and waveforms are scrutinized. Some samples are presented below.

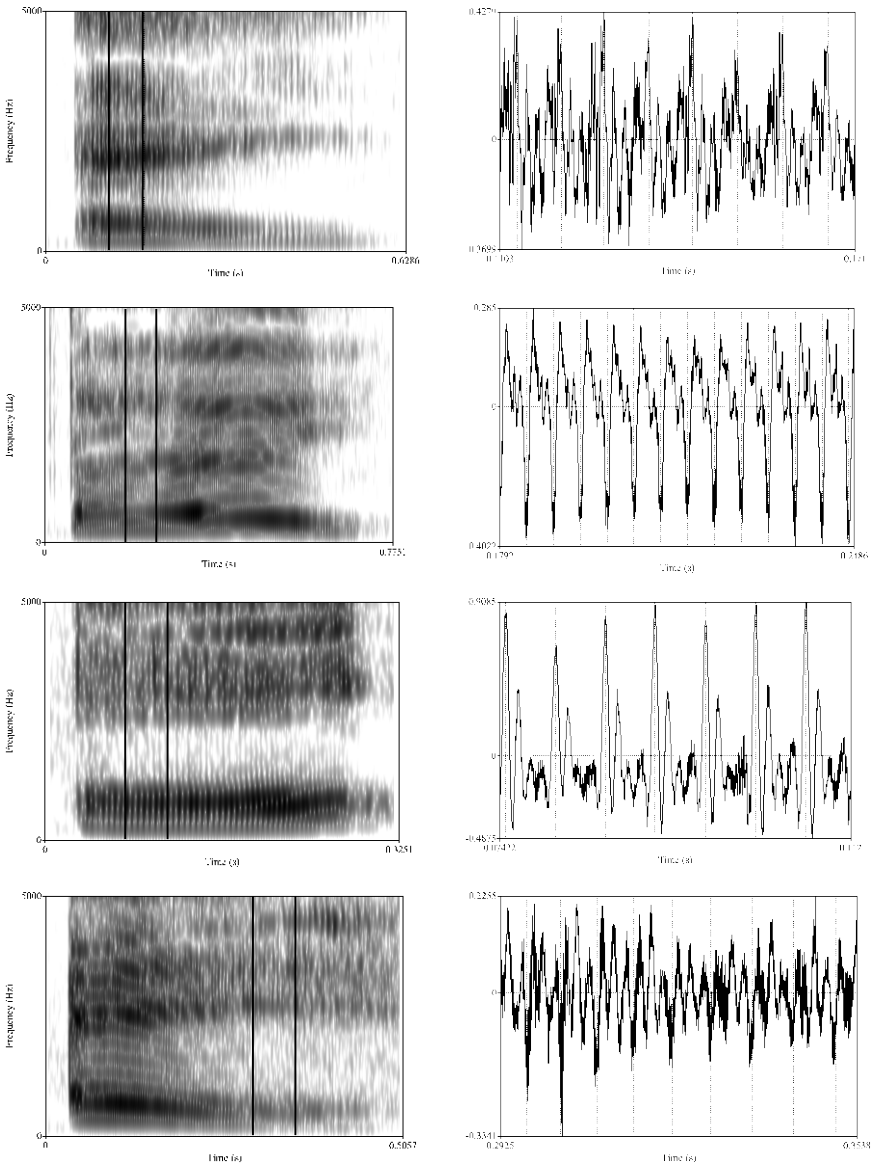


Figure 1. Samples of [ei^{21}] (the first row, by Male 1), [tɵy^{242}] (the second, by Female 3), [ɔʔ^{24}] (the third, by Male 2) and [e^{51}] (the last, by Female 2).

Tones [21 242 24] and the end of [51] are characterized by prickly waveforms and blurry formant structures, suggesting breathy voice. Most tokens of these tones show varying degrees of breathiness. Figure 2 presents the data of HNR35 and H1*-A1*.

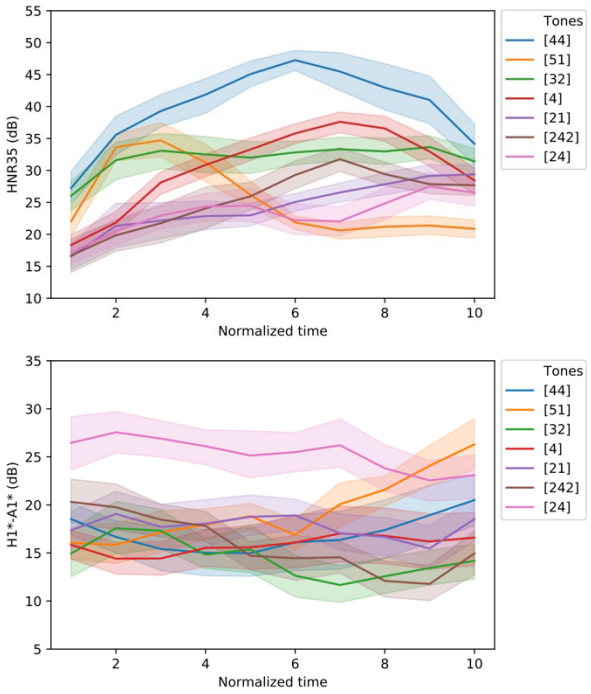


Figure 2. The averages of HNR35 and H1*-A1*, and 95% confidence intervals.

For HNR35, all trajectories start from lower positions, but only [44 51 32 4] rise sharply after the first or second point. Results of repeated measures ANOVA at the 3rd time point show a significant main effect of tones on the values of HNR35 ($F(6, 144)=43.4$, $p<0.001$, $\eta_p^2=0.644$), and post hoc tests (p adjusted by the Holm correction) reveal that the HNR35 values of [21 242 24] are significantly lower than those of other tones ($p<0.001$, except $p=0.005$ for the comparison between [24] and [4]) and no significant difference is found amongst [21 242 24] (all $p=1.000$). Tone [44] has the highest value ($p<0.001$, except $p=0.013$ for the comparison between [44] and [51]). The value of [51] is significantly higher than those of [4 21 242 24] (all $p<0.001$) but no significant difference is found for the comparison between [51] and [32] ($p=1.000$).

Tone [51] shows considerable falling after the 3rd time point and reaches its lowest point after the 6th. The same statistical analysis carried out at the 7th time point also show a significant main effect of tones on the values of HNR35 ($F(6, 144)=86.1$, $p<0.001$, $\eta_p^2=0.782$). Post hoc tests (p adjusted by the Holm

correction) reveal that [51 24] have the lowest HNR35 values ($p<0.001$, except $p=0.004$ for the comparison between [21] and [24]) and no significant difference is found between these two tones ($p=0.496$).

For H1*-A1*, only [24] and the end of [51] stand out. Mean values over the ten time points are obtained for all tones except [51] (which shows large changes across time) and results of repeated measures ANOVA reveal a significant main effect of tones on the values of H1*-A1* ($F(5, 120)=22.5$, $p<0.001$, $\eta_p^2=0.484$) and post hoc tests (p adjusted by the Holm correction) find that only [24] with higher H1*-A1* values is significantly different from other tones (all $p<0.001$). The same statistical analysis at the 9th time point also shows that there is a significant main effect of tones on the values of H1*-A1* ($F(6, 144)=13.9$, $p<0.001$, $\eta_p^2=0.367$), and [51] is significantly higher than all other tones (all $p<0.001$ except $p=0.040$ for the comparison between [44] and [51]) except [24] ($p=1.000$).

Based on the features of waveforms and spectrograms as well as the data of HNR35 and H1*-A1*, it is safe to conclude that [21 242 24] are produced with breathy voice, while [51] is modal-breathy. Measure H1*-A1* can only differentiate [24] and the end of [51] from other tones.

Notes

1. Current affiliation: Department of Linguistics, The University of Hong Kong.
2. In this study, tone letters of Chao (1930) are used to represent tone height. '5' is the highest and '1' is the lowest.
3. Female 1, 2 and Male 1 are born in 1940s. The exact birth years of Female 3 and Male 2 were not noted down, but they are very likely in their sixties and seventies respectively. All recordings were made in 2018.
4. Male 1 produced [tɵy^{32}] as [tui^{32}] (with modal voice). One token of [tɵy^{242}] by Female 2 is replaced by [tɔy^{242}] due to too strong breathiness (aperiodic signal).
5. The data of F0, F1 and F2 are checked against the spectrograms. In a few cases of incorrect tracking (48 out of 1750, about 2.74%), the nearby points closest to the original points are used.

References

Chao, Y. R. 1930. A system of tone letters. Le maître phonétique 45, 24-27.

Donohue, C. 2017. Tones and vowels in Fuzhou revisited. In Segmental Structure and Tone, 99-108. De Gruyter.

Esposito, C. M. 2006. The effects of linguistic experience on the perception of phonation. Doctoral dissertation, University of California, Los Angeles.

Garellek, M. 2012. The timing and sequencing of coarticulated non-modal phonation in English and White Hmong. Journal of Phonetics 40(1), 152-161.

Shue, Y.-L., Keating, P., Vicenik, C., & Yu, K. 2011. VoiceSauce: A program for voice analysis. In Proceedings of the 17th International Congress of Phonetic Sciences, 1846-1849.

Tutoring EFL Students in Portugal: Rethinking writing center methodology

Maria Inês Caldeira de Almeida[1], Sydelle de Souza[2,3]
[1]CEAUL, Universidade de Lisboa, Portugal
[2]Centro de Linguística, Universidade de Lisboa, Portugal
[3]Faculdade de Letras, Universidade de Lisboa, Portugal
https://doi.org/10.36505/ExLing-2020/11/0019/000434

Abstract

The writing center at the University of Lisbon adopts a widely implemented tutoring model—a non-directive approach, focusing on content and organization (higher-order concerns), and grammar and vocabulary (lower-order concerns). This study aims to challenge this model from an English as a Foreign Language (EFL) perspective by reporting and analyzing the results of a qualitative survey of 6 tutors regarding their work at the center. Results show that although all tutors comply with the imported tutoring model, most report difficulties in implementing one or both of its central principles. Therefore, EFL writing centers should harbor a flexible methodology wherein addressed concerns are context-dependent and constantly renegotiated during writing sessions, which could have pedagogical implications in the EFL classroom.

Keywords: EFL, Academic Writing, Writing Center Methodology, Portugal

Introduction

As English has become an essential learning and publication tool in academia (Curry & Lillis, 2004; Graddol, 2006), the popularity of writing centers has increased outside of the United States (Severino & Cogie, 2004). The writing center at the University of Lisbon helps students improve their English academic writing skills, adopting a widely implemented tutoring philosophy based on non-directive methods and a hierarchized dichotomy between higher-order and lower-order concerns.

This study examines and challenges such a model from an English as a Foreign Language (EFL) perspective, considering the tutors and writers' specific linguistic background at a Portuguese university versus American writing centers.

Theoretical background

In the Unites States, writing centers have been shaped by the notions of collaborative learning and writing as both a process and a social activity (Cooper, 1986; Ede, 1989) making these centers places where students, rather than being regarded as passive subjects, can participate in the knowledge building process (Bruffee, 1995). Despite this rejection of "traditional

hierarchies" (Lunsford, 1991), the tutor and writer are not on equal terms, as the latter understands the assignment and its subject, whereas the former possesses knowledge concerning academic discourse and written English (Bruffee, 1995).

In his landmark article on tutoring methodology, Brooks (1991) argues for a "minimalist" approach to tutoring, making the writer "the only active agent in improving the paper." Tutors should help students without direct intervention, avoiding the work of the editor and accusations of plagiarism, while focusing on guiding students through the process of writing. They should not make direct changes to a text, but help students improve as writers, pushing them to find and correct any issues themselves. Thus, as stated by Gillespie and Lerner (2008), an ideal tutorial "would look like two peers having a conversation about writing," where the tutor, instead of directly commenting on a student's work, should prioritize a mostly Socratic questioning method, encouraging the writer to think about his text. Such an approach favored the development of a tutoring procedure that defines higher-order concerns as dealing with anything that is not related to "grammar or word choice" (Gillespie & Lerner, 2008). Therefore, content and organization matters are prioritized in writing center sessions over grammar and correctness, which are considered lower-order concerns.

However, this approach does not take into account the needs of second language writers, who see the writing center "as a place to learn from their tutors both the rhetorical and the linguistic dimensions of writing" (Severino & Cogie, 2004). Thus, following North's (1984) notion that a writing center should define its work "in terms of the writer it serves," it is paramount for each center to consider the linguistic background of both their tutors and tutees when defining its working methodology. An inability to do so has already caused tensions between the non-directive and directive tutoring approaches, as well as frustration among tutors who feel pressured to respect the established hierarchy between higher-order and lower-order concerns (Blau & Hall, 2002; Nicklay, 2012; Severino & Cogie, 2004). Yet, tutoring methodology for EFL writing centers, where both tutor and tutee are second language writers and the linguistic knowledge each brings to the session may be considerably different, remains largely unexplored. Therefore, the aim of the present study is to bring this discussion to the table as the first step towards accomplishing a long-term goal—to construct a methodology suitable for EFL writing centers that considers the linguistic background of writers and tutors alike.

Methodology

This study used Microsoft Forms to create and execute an anonymous online qualitative survey containing Likert-type, multiple choice, and long answer questions. The questionnaire was divided into three sections relating to: (i) the tutors' linguistic background and English proficiency, (ii) the writing center's

tutoring philosophy, i.e., non-directive methods and the hierarchy established between concerns, and (iii) the tutors' overall assessment of their work and experience at the center.

Although the writing center has many tutors[1], only those who volunteered for a minimum of two semesters were asked to take this survey. Thus, six tutors (1F, 5M) with a mean age of 26 years (±2.79) participated in this qualitative survey. Five were native speakers of European Portuguese, and one was a native speaker of German. All self-reported a high proficiency in English (C1-C2 CEFR) across all competencies and learned it pre-puberty (M_{Age}= 9.5 years), i.e., within the critical period of language acquisition.

Results and discussion

Results show that although all tutors comply with the imported tutoring model, most report difficulties in implementing one or both of its central principles. Regarding the hierarchy established between concerns, only one tutor affirms to have frequently addressed higher-order concerns, while half report that the students' level of proficiency in English required them to focus on lower-order ones. Thus, most tutors often felt the need to prioritize syntax-level issues, subverting the hierarchy usually established between concerns. The need for this change can be explained by taking into account the tutees' level of English proficiency, as a high frequency of lexico-grammatical issues in a text can lead to incomprehension, making lower-order concerns more relevant to address than organization or content.

With respect to the non-directive methodology, four tutors state that only a few students would not have benefitted from more direct guidance. Furthermore, although all tutors claimed to have prioritized organization and content issues, they were confronted with grammatical concerns more frequently than they were trained to address, reportedly dealing with lower-order issues in a quarter of all writing conferences. Consequently, the nature of such concerns and the students' linguistic background made it necessary to put aside minimalist tutoring methodology and focus on issues directly. Although such an approach went against traditional writing center philosophy, the students' limited knowledge of the English language could not provide them with enough autonomy to allow for the correction of their own mistakes. As a result, tutors were unable to elicit answers from the writers, making it necessary to switch to a more direct methodology.

Conclusion

Implementing a mostly Socratic methodology, paired with hierarchized and rigid concepts of higher-order and lower-order concerns does not facilitate the tutors' task and could hinder the creation of learning opportunities. Thus, it is necessary to develop a writing center philosophy catering to the English academic writing needs of L1 Portuguese students. Moreover, in EFL writing

centers, mostly non-directive tutoring should give way to a flexible methodology wherein addressed concerns are context-dependent and should be constantly renegotiated during writing conferences.

As this writing center is in its nascent stages, our sample size is limited. Therefore, future work should also consider the experience of EFL tutees while developing similar strategies in linguistically diverse academic contexts. This could also have pedagogical implications in the EFL classroom.

Notes

1. Both the authors of this paper are also tutors at the same writing center. They did not take the survey to avoid any conflict of interest.

References

Blau, S., Hall, J. 2002. Guilt-free Tutoring: Rethinking How We Tutor Non-Native-English-Speaking students. The Writing Center Journal, 23(1), 23–44.

Brooks, J. 1991. Minimalist Tutoring: Making Students Do All the Work. The Writing Lab Newsletter, 15(6), 1–4. https://clas.sa.ucsb.edu/docs/default-source/resources-writing/minimalist-tutoring.pdf?sfvrsn=3e9d2bfb_0

Bruffee, K.A. 1995. Peer tutoring and the "Conversation of Mankind." Landmark Essays on Writing Centers, 87–98.

Cooper, M.M. 1986. The Ecology of Writing. College English, 48(4), 364. https://doi.org/10.2307/377264

Ede, L. 1989. Writing as a Social Process: A Theoretical Foundation. The Writing Center Journal, 9(2), 3–13. https://doi.org/10.1017/cbo9781139173674.015

Gillespie, P., Lerner, N. 2008. The Longman Guide to Peer Tutoring (2nd ed.). Pearson.

Lunsford, A. 1991. Collaboration, Control, and the Idea of a Writing Center. In "The Idea of a Writing Center": Building a Theoretical Foundation (pp. 92–99).

Nicklay, J. 2012. Got Guilt? Consultant Guilt in the Writing Center Community. The Writing Center Journal, 32(1), 14–27.

North, S. 1984. The Idea of A Writing Center. College English, 46(5), 433–446. https://doi.org/10.1017/CBO9781107415324.004

Severino, C., Cogie, J. 2004. Writing centers and second and foreign language writers. Journal of Second Language Writing, 13(3), 165–172. https://doi.org/10.1515/9781614511335-024

The processing of cataphora coreference in Brazilian Portuguese

Pablo Machel Nabot Silva de Almeida[1], José Ferrari Neto[2]
[1]School of Humanities, GENP/LAPROL/GEPROL, Pontifical Catholic University of Rio Grande do Sul, PUCRS, Brazil
[2]Department of Portuguese Language and Linguistics, LAPROL/GEPROL, Federal University of Paraíba, UFPB, Brazil
https://doi.org/10.36505/ExLing-2020/11/0020/000435

Abstract

This study aimed to explore whether the processing of cataphoric coreference in Brazilian Portuguese (BP) is guided by a top-down strategy as postulated by the active search mechanism (ASM) or by a bottom-up routine, as well as if this process is restricted by the Principle C constraint. The results revealed that ASM is not a mechanism used to solve the intended coreference in BP and that the participants have demonstrated sensitiveness to establish the coreference while they had to interpret it. In addition, there is evidence that illicit antecedents are not considered during cataphoric resolution, which suggests that the Principle C constraint impacts on the processing and seems to not be violated during the time course of the computation.

Key words: backwards anaphora, coreference, principle C, active search mechanism.

Introduction

Studies developed on the processing of backwards anaphora have shown that its computation is guided by an active search mechanism, henceforth ASM, which is driven by a top-down strategy without waiting for a bottom-up information to be analyzed. Moreover, the processing of cataphora only proceeds in structural positions that are not ruled out by the Principle C constraint (Chomsky, 1981). This type of computation has been provided by several studies conducted in different languages such as English (Kazanina et al., 2007). Additionally, concerning the impact of the Principle C constraint in BP, the results that have been achieved are incipient and inconclusive (Lessa & Maia, 2011; Maia, Garcia & Oliveira, 2012).

Therefore, we raised the question whether the parser of a pro-drop partial null-subject language like Brazilian Portuguese (BP), which is in a process of parametric changing, would act in the same way as the parser of previously studied languages (see Barbosa, Duarte & Kato, 2005). Hence, we aimed at investigating whether the processing of cataphoric coreference in BP is actually guided by the ASM or by a distinctive routine less active and if it is constrained by the action of the Principle C constraint.

Methodology

In order to clarify these open questions, we conducted a non-cumulative self-paced reading experiment that was run in the Language Processing Laboratory (LAPROL). Participants were 71 native speakers of BP from the Federal University of Paraíba (UFPB), and all of them gave informed consent and volunteered to participate. We employed the gender-mismatch paradigm (Van Gompel & Liversedge, 2003) to conduct the experiment that was designed using the software Paradigm, version 2.5.0, in a laptop Acer Intel Core I5-2450M, 2,5 GHz, 4 GB DDR3. It had five conditions, four of which were organized in a 2 X 2 Latin Square design with the factors *constraint* (Principle C vs. No-Constraint) and *gender congruency* (Match vs. Mismatch between the pronoun and the subject of the second clause). An additional fifth control condition characterized by an anaphoric coreference was added. In each trial of the experiment an interpretative question was included to check if the participant identified if the coreference was established or not. A sample set of experimental items can be seen in example (1):

1. a. No-Constraint Match/No-Constraint Mismatch/Name Control
   Visto que tempos atrás enquanto ela$_i$ (Carla$_i$) estava estudando para a seleção *Carmen$_i$ (Paulo)* trabalhava o dia inteiro, Paulo quase não a via para namorar por ela viver muito ocupada//Carla$_i$ (ela$_i$) prontificou-se a vê-lo para namorar nas horas vagas disponíveis para os dois. (a sample of the stimuli presented in BP).
   Since long ago while she$_i$ (Carla$_i$) was studying for the exam *Carmen$_i$ (Paulo)* worked all day, Paulo rarely saw her for dating because she was very busy//Carla$_i$ (she$_i$) volunteered to see him for a date in the spare time available to them.

   b. Principle-C Match/Principle-C Mismatch
   Visto que tempos atrás ela$_i$ (Carla$_i$) estava estudando para a seleção enquanto *Carmen (Paulo)* trabalhava o dia inteiro, Carla$_i$ (ela$_i$) se sentia constrangida por não contribuir também com as despesas. (a sample of the stimuli presented in BP).
   Since long ago she$_i$ (Carla$_i$) was studying for the exam while *Carmen (Paulo)* worked all day, Carla$_i$ (she$_i$) felt embarrassed for not also contributing to expenses.

For the analysis of the reading-time data, the critical segment concerning the first NP that was taken as the first potential *postcedent*[1] was analyzed. Also, the interpretative questions were analyzed as well. The R lme4 package (Bates, Maechler, Bolker & Walker, 2015) was used for the analysis and linear mixed-effects regressions were calculated for each condition of interest.

Results and discussion

After removing outliers from the data, which resulted in an approximately normal distribution of the data, a linear mixed-effect model was used, adopting the reaction times (RT) from the reading of the critical segment as dependent variable and both subjects and items as random effects. *Type of principle*

(Principle-C and No-Constraint) and *Matching* (Match and Mismatch) were considered as independent variables. The results of the statistical analysis are shown in Table 1:

Table 1. Results of the linear mixed-effect model statistics.

	Estimate	Std.Error	df	t value	Pr(>|t|)
(Intercept)	970.56	35.84	81.17	27.079	< 2e-16*
PrincipleCNoConstraint	-142.54	25.60	181.14	-5.569	9.15e-
MatchMismatch	-41.67	23.02	818.11	-1.811	08*
PrincipleCNoConstraint : MatchMismatch	40.38	31.97	1444.15	1.263	0.0706 0.2067

The lack of a gender-mismatch effect (GMME) emerged in the pairwise comparison between the No-constraint conditions, suggesting that the processing of backwards anaphora in BP is not driven by an active and top-down routine resembling the ASM as predicted ($t(70) = 35.24$, $p<0.5$). Furthermore, the lack of significant differences between Match and Mismatch conditions reveal that the phi-feature of gender does not influence that kind of processing regardless of the active presence of the Principle C constraint. On the other hand, the significant differences revealed in the comparisons made between Principle C and No-Constraint conditions show that the Principle C constraint interferes in the processing of cataphoric coreference as well as longer reading times observed in the No-Constraint conditions illustrate that the processing of cataphora is harder when there is no Principle-C constraint operating in the referred coreferential processing.

The responses given by the subjects to the interpretation of the coreference, here also considered as a dependent variable, were analyzed by using a chi-square test (X^2). The results are shown in Tables 2 and 3 below:

Table 2. Results of the chi-square test (X^2): gender.

	Match	Mismatch	Total
Expected	514	467	981
Unexpected	338	385	723
Total	852	852	1704

$X^2 = 5{,}08$, D.F = 1, $p<0.02$.

Table 3. Results of the chi-square test (X^2): constraint.

	Principle-C	No-Constraint	Total
Expected	444	537	981
Unexpected	408	315	723
Total	852	852	1704

$X^2 = 20{,}335$, D.F = 1, $p<0.05$.

In relation to the interpretation of cataphoric coreference, the results point out that BP speakers find it easier to correctly establish the coreference when there is no constraint, whereas they hesitate to reject the coreference and establish the appropriate disjoint reference when the Principle C constraint operates. Finally, although in both Match and Mismatch conditions BP speakers have demonstrated to be able to correctly establish the expected coreference, the results demonstrate that the answers were more consistent in the Match conditions. This pattern signals the phi-feature of gender is also able to modulate the interpretation of cataphoric coreference, even though it is not able to influence it in the earlier processing phase. In sum, this suggests that the morphological congruence makes it easier for BP speakers to accept the suitable coreference as opposed to the morphological incongruence that makes harder such a coreferential interpretation towards the proper disjoint reference.

Notes

1. See Almeida (2016) for the original use of the term *poscedente* (postcedent) in BP.

Acknowledgements

We thank to CAPES – Financing Code 001 and to CNPq – Project 870599/1997 6, concerning Process 140275/2019 3. We also are grateful to Lilian Cristine Hübner for her revision and constructive comments.

References

Almeida, P. M. N. S. D. 2016. O Processamento da correferência catafórica pronominal pessoal em Português Brasileiro. Master's thesis, Federal University of Paraíba – UFPB, João Pessoa.

Barbosa, P., Duarte, M. E. L., Kato, M. A. 2005. Null subjects in European and Brazilian Portuguese. Journal of Portuguese Linguistics 4(2), 11-52.

Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67(1), 1-48.

Chomsky, N. 1981. Lectures on Government and Binding. Dordrecht, Foris.

Kazanina, N., Lau, E. F., Lieberman, M., Yoshida, M., Phillips, C. 2007. The effect of syntactic constraints on the processing of backwards anaphora. Journal of Memory and Language 56(3), 384-409.

Lessa, P. D. C., Maia, M. 2011. O processamento da correferência de expressões-R com restrições lexicais em Português Brasileiro. In Brazilian Linguistics Association (ed.), Proc. 7th Abralin International Congress, 3598-3606, Curitiba, Brazil.

Maia, M., Garcia, D. C. D., Oliveira, C. D. 2012. The processing of conceptual anaphors and fully specified pronouns in intra-sentential contexts in Brazilian Portuguese. ReVel: Revista Virtual de Estudos da Linguagem 10(6), 200-219.

Van Gompel, R. P. G., Liversedge, S. P. 2003. The influence of morphological information on cataphoric pronoun assignment. Journal of Experimental Psychology: Learning, Memory, and Cognition 29(1), 128-139.

The grammar behind word association tasks

Sydelle de Souza[1,2], Alina Villalva[1,2], Carina Pinto[2,3,4]
[1]Faculdade de Letras, Universidade de Lisboa, Portugal
[2]Centro de Linguística, Universidade de Lisboa, Portugal
[3]Escola Superior de Saúde, Politécnico da Leiria, Portugal
[4]Centro de Inovação em Tecnologias e Cuidados de Saúde, Portugal
https://doi.org/10.36505/ExLing-2020/11/0021/000436

Abstract

Most word association tasks (WATs) focus on semantic representations but can also illuminate other linguistic phenomena (Van Rensbergen et al., 2015). This study reports a WAT comprising 152 Portuguese verbs and their corresponding compositional deverbal action nouns ending in *-ção*, that aims to test whether the morphological nature of the stimulus conditions the response. The results suggest that (i) the morphological structure of the stimuli does condition the response; (ii) associations can be constrained by grammatical factors other than semantics; and (iii) L1 Portuguese speakers are more likely to respond with a word belonging to the same word family if the stimulus is a complex word.

Keywords: morphological processing, Portuguese, word association task, lexical semantics, complex words

Introduction and theoretical background

The first known word association task (WAT) was performed by Sir Francis Galton in 1883 (Galton, 1907) and has since been frequently used by experimental and differential psychologists. Later, WATs became instrumental in experimental linguistics and research pertaining to the mental lexicon (Clark, 1970). They have been used to study semantic and lexical access in a variety of areas ranging from child language acquisition (Sheng et al., 2014) and second language acquisition (Nissen & Henriksen, 2006; Zhang & Koda, 2017) to clinical studies involving dementia (Eustache et al., 1990) and deep dyslexia (Colangelo et al., 2003). Although most WATs only focus on semantic representations, it has been argued that data obtained from them "often shed light on other linguistic phenomena as well" (Van Rensbergen et al., 2015).

The instructions—read the following words and write the first word that comes to your mind—raise doubts about whether the response is truly the first word that comes to the subject's mind. These doubts along with the emergence of other sophisticated methodologies led to the decline of WATs. However, a recent study by Playfoot et al. (2018) claiming that "word association responses are likely to reflect the first word that participants activate in their lexicon", may justify a revival.

Therefore, the present study aims to determine (a) if WATs are constrained by grammatical factors other than semantics, (b) if the morphological structure of the stimuli triggers specific types of responses, and (c) if L1 Portuguese speakers favor certain lexical activation pathways for complex words.

Methodology

A WAT comprising 152 Portuguese verbs (e.g. *aceitar* 'to accept') and their corresponding deverbal action nouns, formed by the suffix *-ção* (e.g. *aceitação* 'acceptance') was constructed. Frequency values were considered (*Corpus de Referência do Português Contemporâneo*) and the stimuli (2-6 syllables) were divided into: high (>1000 tokens), medium (101-1000 tokens), and low frequency (<100 tokens) words.

The subject sample consisted of 22 proficient native speakers of European Portuguese (EP). All were university students (M_{Age} 20 years ±1.95) and presented no language pathologies. The task, executed individually through Google Forms, was completed in approximately 90 minutes. All participants were present in the same room and were instructed to enter the first word that occurred to them on reading the stimuli.

6688 responses were collected and manually tagged based on their relationship to the stimulus, according to the following categorization:

A. **Strictly semantic relationship:** Synonyms, hyponyms, etc. (*acusar* 'to accuse' – *culpar* 'to blame')
B. **Morphosemantic relationship**: Derivative of stimulus (*ligar* 'to connect' – *ligação* 'connection'), stimulus base word (*aceitação* 'acceptance'– *aceitar* 'to accept'), parallel derivatives of the base word (*admiração* 'admiration' – *admirável* 'admirable')
C. **Syntagmatic relationship:** Complement or modifier of the stimulus (*ligar* 'to connect' – *telemóvel* 'cell phone')
D. **Other:** Possible phonetic resemblance, grammatically random (e.g. *polinizar* 'to pollinate' – *polir* 'to polish')
E. **Invalid:** Response is not a word
F. No response

Results and discussion

The results show that only 6.3% of the total responses are random, invalid, or no responses while all other responses are grammatically linked to the stimulus. Global results (N + V) show that 46.3% of the responses establish a strictly semantic relationship and 33.9% a morphosemantic relationship. Syntagmatic relationships were established in 13.5% of the responses. Disaggregating the global results by word-class (N vs. V) shows a noticeable

stability for the syntagmatic relationship (N = 13.6%, V = 13.4%) and a deviation close to 5% in the other two cases (Table 1).

Table 1. Global Results N vs. V.

	Semantic Relationship (A)	Morphosemantic Relationship (B)	Syntagmatic Relationship (C)	Other (D+E+F)
Nouns	43.8%	36.2%	13.6%	6.4%
Verbs	48.8%	31.5%	13.4%	6.3%

The semantic pathway percentage is higher for verbs, and the morphological pathway is higher for nouns, confirming the hypothesis that the morphological nature of the stimulus triggers certain types of responses. A further disaggregation of the results by morphological complexity (simplex V vs. complex V and N) shows that complex words elicit increased morphologically related responses, i.e., morphological activation is clearly triggered by morphologically complex stimuli (Fig. 2).

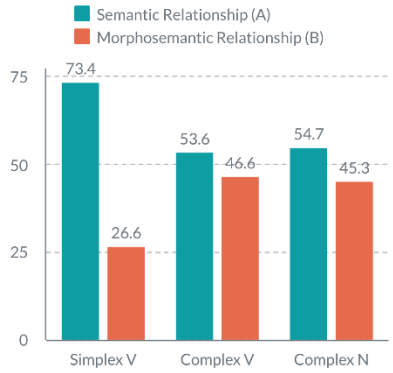


Figure 2. Simplex V vs. complex V and N.

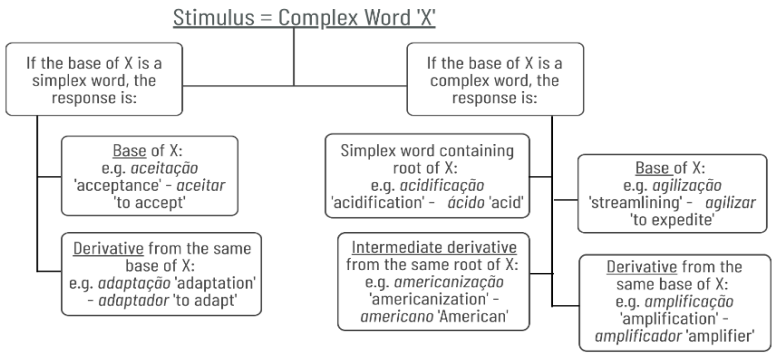


Figure 3. Typology of responses.

Additionally, simplex words containing the embedded root appear to be the most popular choice when a morphological response is given. Therefore, irrespective of the degree of complexity of the derivative, the embedded lexeme seems to be successfully identified (Fig. 3).

Conclusion

In sum, the results suggest that the morphological structure of the stimuli is of paramount importance to the response and that word associations are constrained by grammatical factors other than semantics. Finally, for complex stimuli, L1 Portuguese speakers are more likely to respond with a word belonging to the same word family. As this sample is limited to deverbal action nouns ending in *-ção*, the findings cannot be extrapolated to other complex words without replicating this test. Other tests have been performed with complex words ending in *-vel* and *-dor*. Their results are currently under analysis.

References

Clark, H. H. 1970. Word Associations and Linguistic Theory. New Horizons in Linguistics, 1–8.

Colangelo, A., Stephenson, K., Westbury, C., Buchanan, L. 2003. Word associations in deep dyslexia. Brain and Cognition, 53(2), 166–170.

Eustache, F., Cox, C., Brandt, J., Bernard, L., Pons, L. 1990. Word-Association Responses and Severity of Dementia in Alzheimer Disease. Psychological Reports, 66(4), 1315–1322.

Galton, F. 1907. Inquiries into Human Faculty and its Development (Everyman S). J. M. Dent & Co. https://doi.org/10.16309/j.cnki.issn.1007-1776.2003.03.004

Nissen, H. B., Henriksen, B. 2006. Word class influence on word association test results. International Journal of Applied Linguistics, 16(3), 389–408.

Playfoot, D., Balint, T., Pandya, V., Parkes, A., Peters, M., Richards, S. 2018. Are word association responses really the first words that come to mind? Applied Linguistics, 39(5), 607–624.

Sheng, L., Mcgregor, K. K., Marian, V. 2014. Lexical – Semantic Organization in Bilingual Children : Evidence From a Repeated Word Association Task. Journal of Speech Language and Hearing Research, 49(June 2006), 572–587.

Van Rensbergen, B., Storms, G., De Deyne, S. 2015. Examining assortativity in the mental lexicon: Evidence from word associations. Psychonomic Bulletin and Review, 22(6), 1717–1724.

Zhang, D., Koda, K. 2017. Assessing L2 vocabulary depth with word associates format tests: issues, findings, and suggestions. Asian-Pacific Journal of Second and Foreign Language Education, 2(1), 1–30.

Complements vs. adjuncts and mouse-controlled reading

Armine Garibyan
Friedrich-Alexander University Erlangen-Nürnberg, Germany
https://doi.org/10.36505/ExLing-2020/11/0022/000437

Abstract

The relationship between sentence processing and cognitive demand has received a lot of attention in the past decades. In valency theory, some elements of the sentence are determined by the verbs either in terms of their form or by their presence (Herbst & Schüller 2008). It has to be said that little attention has been paid to the processing of such fundamental categories in the theory of syntax. On the one hand, this is remarkable since given the amount of research, we still do not know whether this distinction is psychologically real, or whether it only serves a lexicographic and pedagogical purpose. On the other hand, there is a consensus among linguists about the problematic character of the distinction itself even on a more theoretical level (Dowty 2000; Herbst & Schüller 2008). Therefore, this study attempts to explore whether complements and adjuncts are associated with different kinds of processing. To answer the research questions, an experiment consisting in a mouse-controlled reading task has been designed. To the best of our knowledge, this is a new method in psycholinguistic research. The paper presents the results of a pilot study.

Keywords: valency, complements, adjuncts, processing, mouse-controlled reading

Introduction

In this paper, it has been attempted to uncover whether such theoretical concepts as complements and adjuncts are psycholinguistically valid. The question is topical given the problematic character of such a distinction even on a theoretical level (Hoffmann 2007; Herbst & Schüller 2008).

A popular and reasonable way to do so would be using an eye-tracker to create a natural setting for reading with items embedded in sentences. However, the costs for eye-tracking experiments are vast, which involves the price of the equipment, laboratory costs, research assistants, etc. In this study, it has been attempted to create a method which would resemble the natural reading process, however with little or no cost – the mouse-controlled reading (MCR).

It is expected that in comparison with adjuncts, complements will present processing advantage given their predictability from the preceding verb. Also, it is believed that those participants with higher level of language proficiency will benefit from this difference the most, given their higher familiarity with these item-specific patterns.

Experimental setting

Materials and methods

The test battery included a personal background questionnaire, a c-test as well as a computer-based experiment (MCR). The c-test was administered as an economical measure of general language proficiency (Klein-Braley 1985). The MCR method consisted in one sentence at a time presented on the screen where the text was always hidden behind (or masked by) a white foreground. By moving the mouse, the participant could see the text behind the mouse[1,2] (see Figure 1). By limiting the visible region to a small circle, it was possible to tell where exactly the participant was looking at any given point in time. Backtracking was possible, however in terms of one trial only.

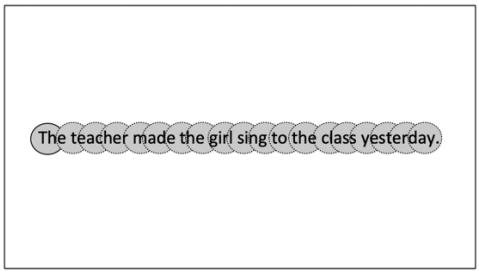


Figure 1. An example sentence of the MCR task. The sentences written on grey background were hidden behind a white foreground. As soon as the reader proceeds from left to right, the previous spot gets covered again.

Participants were asked to read and decide whether the sentence was grammatical or not. Altogether, there were 55 sentences: 15 containing complements, 15 adjuncts and 25 (ungrammatical) fillers for the distractor task. All complement and adjunct units were controlled for length and frequency in the pair and counterbalanced to avoid priming. In (a) and (b) below, the target units are underlined. The verb always appeared in the same word form.

(a) The mother went to search <u>for cigarettes</u> in her son's room.
(b) We can only search <u>with permission</u> on private land.

Being a pilot study, ten participants with L1 German were recruited. The MCR task was presented using the software *PsychoPy* (version 2020.2.1) on a MacBook Pro 13. Participants' mouse movements were recorded at a sampling rate of 60Hz (approximately every 16.6 msec), i.e. x and y coordinates as well as the time spent at each point. Statistical analysis was carried out in *R* programming language (R Core Team 2019).

Results

A linear mixed-effects model was used to account for individual differences in participants as well as the differences between items. Only target regions (complements and adjuncts) were analysed.

A first inspection of the density plots (Figure 2) shows that participants spent more time reading complements than adjuncts[3]. Although the mean time for reading complement units is 1.92 sec and for adjuncts 2.20 sec, statistical analysis has not detected any significant effects (p = .322). Neither has the interaction between condition and language proficiency reached a significance level. However, the analysis showed that responses to the distractor task (i.e. grammatical vs. ungrammatical) were significantly associated with reading times (p = .002), so higher number of correct answers resulted in shorter reading times.

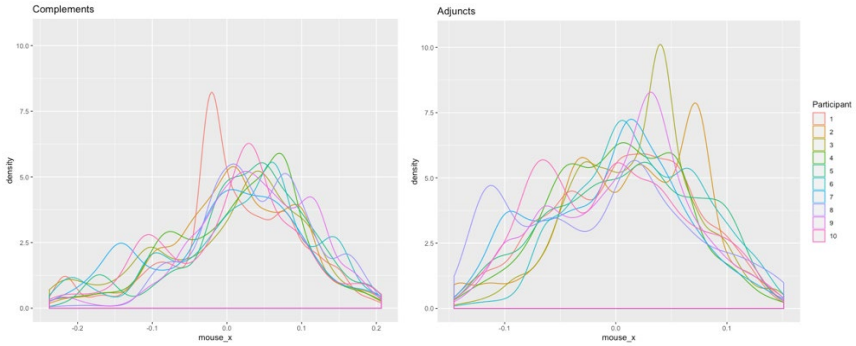


Figure 2. Density plots showing the distribution of data points while reading complement vs. adjunct units by participant.

The findings of the study appear to be in contrast with established literature in the field according to which the predicator entails information on both the argument frame as well as the thematic roles (Rickheit & Sichelschmidt 2007). What comes next is the reasoning on why the study needs to be continued.

The inspection of Figure 2 allows for further interpretations. First of all, it is clearly visible that there is more alignment between participants' reading flow of the complement units rather than adjuncts, which proves the presence of an unbiased (formal) component predicted by the verb. Second, the general consistency of the reading procedure (in both plots) allows us to think that the MCR should be further studied to explore its potential and validity. Several measures will be taken to continue the study. First, more participants will be tested to reach statistical power. Second, to check whether the absence of the effect on the main condition was caused by the categories themselves or the

method, a different method will be used to verify the effect. Finally, a different phenomenon will be used in attempt to verify the validity of the MCR method.

Notes

1. Imagine that you are trying to read a letter in a dark room holding a candle. At each point in time, you can only read the part of the letter at which the candle is pointing.
2. A mouse was chosen over a trackpad to allow for free hand movement (not limited by the size of the trackpad) and thus, a more natural flow of movement.
3. Every point x which the participant passed was associated with 16.6 msec (the refresh rate). The plot shows the frequency of these passings.

Acknowledgements

I would like to thank Thomas Herbst, Ewa Dabrowska, Peter Uhrig and Laura Becker for their valuable feedback on different parts of the study.

References

Dowty, D. 2000. The dual analysis of adjuncts/complements in categorial grammar. ZAS Papers in Linguistics 17. 53-78.

Hoffmann, Th. 2007. Complements versus adjuncts: A construction grammar approach of English prepositional phrases. Occasional Papers in Language and Linguistics (University of Nairobi) 3. 92-119.

Herbst, Th. & Schüller, S. 2008. Introduction to syntactic analysis: A valency approach. Tübingen, Gunter Narr Verlag.

Klein-Braley, C. 1985. A cloze-up on the C-Test: a study in the construct validation of authentic tests. Language Testing 2(1). 76-104.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rickheit, G. & Sichelschmidt, L. 2007. Valency and cognition – a notion in transition. In Herbst, Th., Götz-Voteller, K. (eds.), 2007, Valency: theoretical, descriptive and cognitive issues, 163-182. Berlin/New York: Mouton de Gruyter.

Sound dimensions and formants

Grandon Goertz[1], Terese Anderson[2]

[1]Department of Language, Literacy and Sociocultural Studies, University of New Mexico, US

[2]Department of Chemistry and Chemical Biology University of New Mexico, US

https://doi.org/10.36505/ExLing-2020/11/0023/000438

Abstract

Every physical event that can be observed can be measured and described, including sounds. This paper discusses computer algorithms that were developed to depict vowels and speech sounds in their three dimensions: frequency, energy, and time. Each vowel has a separate distinguishable shape based on its dimensions. Two-dimensional vowel plots can be more accurately represented in three-dimensional plots. Algorithms using the Chebyshev Transform were written and vowel speech signals were converted to accurate numerical data sets that were examined and then plotted. Comparisons of vowels can be made, based on their sonic shape. This algorithm also used the Singular Value Decomposition (SVD) to measure, vowel formants giving clear formant regions with the frequency regions identified on the y-axis plots.

Keywords: Three-dimensional depiction, Chebyshev Transform, formant

Introduction

Sound is a vector-defined, moving, measurable physical quality that can be depicted in three dimensions, as illustrated by Khutoryansky (2019).

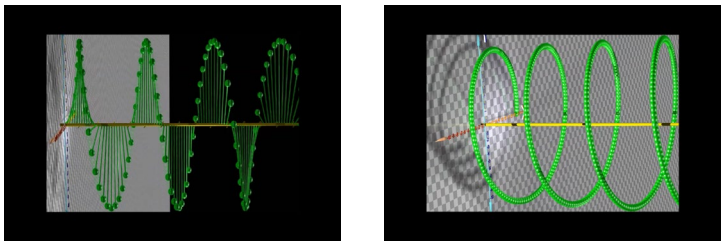


Figure 1. Side view of a pure tone, showing vectors. The right picture shows the same sound at an angle, illustrating the three dimensions.

Algorithms were written to measure voice data in order to produce accurate depictions of sounds by first converting the data to an accurate numerical base, then analyzing the data with a Chebyshev Transform (Boyd 2001), and finally using the Singular Value Decomposition (Gold, Morgan 2001). The next step was to produce three-dimensional plots of the vowels and measured plots of the formants (Johnson 2012). Empirical data can be evaluated, and the data

used to create a picture for each vowel that depicts the waveform envelope and shows the sound shapes with its properties of frequency, time, and energy.

Method

Formant values produced with a 2-meter refractometer, a mechanical device which separates sound into its components, did not corelate to the calculated formants from a popular computer program used by linguists. It was also observed that formants produced by a traditional FFT-based program would not plot in a way that is consistent with the understanding of formants as being harmonic-frequency intense locations. We experimented with several sound files of various languages from the University of California phonetics website (UCLA: phonetics.ucla.edu) and vowel sounds from the first author. It was discovered that sound data was not correlating accurately when it was plotted using base 10. Base 10 is also known as the 0-9 decimal scale.

We assessed if the data was better represented in a different numeric base. Evaluations were done with a mathematical procedure, Principal Component Analysis (PCA), to determine which numeric base described the data best. The process of determining the base began with evaluating each vowel using its T-square values and histograms. The data represented the correlation of the frequency values of the formants of German, English, Swedish, French, and Japanese vowels.

Base e, which occurs very often in biological phenomenon such as growth, was examined and found to be the most accurate base to represent sound frequencies and propagation. Base e, also called the natural logarithmic space, uses a numerical scale in which the numbers ascend in a logarithmical sequence. When base 10 is used the numbers are misrepresented as being too small.

In figure 2 below, the top grey line represents the numerical values of base 2, and the red line just under it represents base e (the value of base e is 2.71828). Moving toward the x-axis the grey lines represent bases 3, 4, 5, 6,7, 8, and 9. The bottom green line represents base 10, a log scale, the scale that is most familiar.

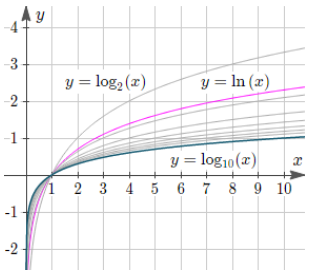


Figure 2. Bases. The base used for sound is represented by the red line.

An algorithm written in Matlab, processes the speech data by first converting the .wav or .AIFF file into frequencies. The frequency data was transformed using Chebfun (Trefethen 2000). A time vector was computed, and the frequency axis was scaled according to the data size. The transformed frequency data was then converted to base e, the natural log values. Energy was also computed for each data point.

The SVD was next computed, the values were checked for accuracy, and the result is a matrix of formants. SVD is an empirical based matrix computation that effectively reduces empirical data to the most significant data values that bests represents the original data. This allows for detection of the true structure of data for modeling.

Results and discussion

The first discovery is the ability to visualize the distinct vowel shapes as they propagate through space (without dispersion) in three dimensions. After converting the sound data to base e, it was possible to plot the data as a picture of a three-dimensional event.

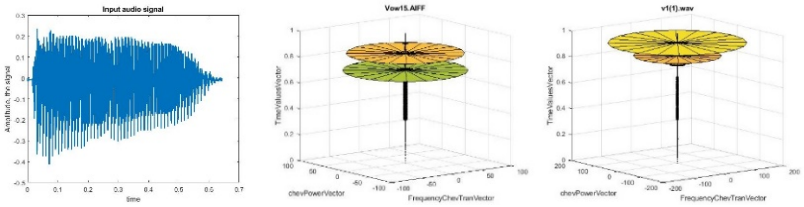


Figure 3. A two-dimension waveform of /o/ produced by a male speaker. Center is the three-dimensional depiction of the same /o/. On the right is the vowel /i/, also by a male speaker. (Black center lines in the center and right pictures are computer-generated zero values.)

Using the data in the correct natural log e space allowed us to also plot formants that were un-ambiguous. The SVD calculation method produced verifiable frequency values that, when plotted, show clear formant banding.

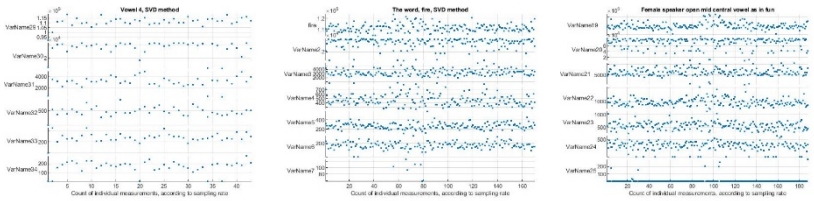


Figure 4. The vowel /ʉ/ on left, and center, the word English word 'fire' spoken by a male. Both plots show clear and calculated formant bands. On the right, /ʌ/ spoken by a female, showing formants.

Formants exist in clear bands, typically not more than seven. The frequency zones are shown on the y-axis and time is indicated by sample number on the x-axis. Both formants and three-dimension depictions define the vowel, in natural log space.

References

Johnson, K. 2012. Acoustics and Auditory Phonetics. MA, Wiley-Blackwell Malden.

Boyd, J. 2001. Chebyshev and Fourier Spectral Methods. Mineola, NY, Dover Publications, Inc.

Gold, B., Morgan, N. 2000. Signal Processing and Perception of Speech and Music. NY, John *Wiley & Sons*.

https:// www.youtube.com/user/EugeneKhutoryansky. January 25, 2019.

Trefethen, L. 2019., *Spectral Methods in MATLAB*, SIAM, 2000.

The meaning of democracy vs. ideology

Agnieszka Grażul-Luft
Faculty of Humanities and Social Sciences, Mazovian State University in Płock, Poland
https://doi.org/10.36505/ExLing-2020/11/0024/000439

Abstract

The subject of research refers to changes in the meaning of the lexeme democracy over the past 20 years, noticeable in texts from the Polish press of specific ideological profile. The authors of press texts often expand the meaning, saturating it with emotions and evaluating. On the basis of examples of using a word in texts, there have been definitional sentences created testifying to the extensions of meanings comparing to those found in the Polish language presented in dictionaries. The analysis demonstrates that the understanding of the word democracy depends on the political context and is subject to modifications resulting from ideological entanglements.

Keywords: ideology, meaning, lexeme, democracy, political discourse

Introduction

Ideology, influencing the shaping of public life and emotional attitudes towards politics, does influence the language. The authors of press texts often expand the meaning, saturating it with emotions and evaluations. The use of language is influenced by the context, which leads to extensions and changes of meaning (e.g. Cruse 2001, Yule 1994, 1996, Brown 1998). The aim of the research was to analyze the differentiation of the meaning of the lexeme democracy in the political discourse on the basis of contextual uses in writings classified as right-wing and left-wing, so as to answer the question whether and to what extent ideology influences the shaping of meaning.

The analysis of the meanings of lexemes-concepts has been focused around the issue of the dyad, in this case created from the antithetical concepts of the right- the left. Although the dichotomy is the tool right for describing the socio-political realities of the turn of the century, in the case of the following research it is not the most important what substantive content the two concepts entail. What is important is that they constitute the two opposing sides of the political conflict. The opposition of those categories represents the typical dyadic mindset (Bobbio 1996). Binary nature has become the basis for the selection of press titles and the very description of changes in meaning. Even though the political situation in Poland in the second decade of the 21st century has changed so much that the opposition between the right-wing and the left-wing has been relegated to the background while the opposing sides of the political conflict created two parties located within the right-wing and the center, the dichotomous context has not lost any of its topicality.

Methodology

The study was conducted in three stages. The first main stage covered the analysis of press texts of 1999-2001, i.e. a decade after the political transformation in Poland. The extensions of the lexeme democracy observed and the conclusions of the main analysis were then confronted with the press material of 2011-2013 (stage 2) and 2017-2019 (stage 3) in order to answer the question regarding the degree of consolidation of the phenomena observed in the Polish language.

The research in the first stage covered six Polish newspapers and magazines: three identified as right-wing and three as left-wing ones. There were excerpts made in journalistic texts to gather examples of using the lexeme democracy. The analysis of changes and semantic extensions was preceded by a review of definitions from several dictionaries of the Polish language. In the description of changes and semantic extensions, the basic unit of definition are not isolated features, but sentences. That type of method allowed the users of the language to understand the lexeme from their subjective point of view. In the second and third stages, there were texts from four (two right-wing and two left-wing) of the six originally selected press titles analyzed as the other two have disappeared from the press market.

As a result of the excerption in the first stage, there were 179 examples of the use of the word democracy collected (in the form of quotations). The frequency of the lexeme usage varied in individual press titles (e.g. due to the fact that they included dailies, weeklies and a monthly). In the research of stages II and III it was assumed that 100-120 citations containing examples of using the democracy lexeme would be collected to achieve the goal assumed. The material was analyzed in terms of the topicality of the changes and semantic extensions described in the implementation of the first stage.

Results

Basing on the analysis of the contextual use of the lexeme democracy, there were defining elements created absent from dictionary definitions. The meaning of the lexeme in contexts is evaluated and is to a significant degree subject to ideological manipulation.

The key to understanding the lexeme democracy seems to be the division between the ideal and the reality (Dahl 1995). In the material analyzed that specific duality is the basis for thinking about democracy. On one hand, it is the ideal which in Poland was hard to achieve, for which we fought. On the other hand, the ordinary everyday life of our democracy leaves much to be desired.

Both in texts of right-wing and left-wing magazines it is emphasized that democracy is the system with certain values and principles binding. They are juxtaposed with the opinions on actual democracy that does not meet the expectations. On the basis of the texts analyzed, real democracy can be described by means of definitional sentences.

'Democracy is a system caring just for the interests of the ruling class'
Although democracy is assumed to be the power of the people, when the representatives who exercise power are elected, it ceases to be the system serving the general public. What then becomes of importance is the wellbeing of those in power often achieved at the expense of the society. The feature is considered independent of the political profile of the ruling ones.

Democracy is the system beneficial only to those who managed to get rich'
This explication, like the one presented above, is based on the opposition, the dyad. In the former sense, the contrast concerned those in power and the rest of the citizens. Here the opposition is formed between the capitalists (all those who have managed to get rich under the new system) and the rest of society.

'Democracy is a system whose disadvantage is that all citizens have the right to vote'
There is a paradox in this sense. The freedom of choice and the exercise of power by all citizens thanks to voting is a kind of cornerstone of democracy, which the society is highly proud of. However, it has become a curse. Democratic practice seems far from ideal, and in the opinion of the writers not all citizens possess the appropriate predispositions to make responsible choices. Another problem is that the right to vote by all citizens does not guarantee that the elected government or the adopted law is in line with the will of the majority (which is the essence of democracy), as too few people participate in the elections.

'Democracy is a system under threat'
The texts analyzed show that there are numerous dangers to democracy, including, for example, internal political conflicts.

'Democracy poses a threat'
The sense of threat results from the basic principles of the functioning of democracy, it seems obvious to the writers, however, they do not specify what to fear.
The semantic extensions presented found their confirmations in the texts analyzed in stages II and III.

Conclusions

According to press analyses, Polish democracy turns out to be actual anti-democracy. In the material gathered, there are few statements indicating that Polish citizens are satisfied with the shape of democracy. Citizens are happy that there is democracy in Poland at all, but are still worried regarding its quality.

The view that there is significant discrepancy between ideal and real democracy unites all sides of the political scene. They indicate the same negative features and accompanying phenomena. Still, it is clear that the assessment of the functioning of democracy in Poland depends on which side of the political conflict is in power. The state of democracy in Poland is assessed higher when the authors of the texts are on the same side of the political conflict as the ruling party. After the power shift, there are changes in the attitude of the writers noticeable immediately.

At the same time, each side displays a certain more or less precise vision of ideal democracy. In that matter, there are ideological divisions coming to life and imaginary democracies taking different shapes depending on their views. The left-wing contests, for example, the xenophobic right-wing movement associated with extreme nationalism. The right wing, on the other hand, is far, in its perception of a democratic system, from, for example, the extreme left-wing connection of democracy with socialism.

In the press basing on ideological message, lexemes become a tool of persuasion. With the use of conversational implicatures, the expression democracy, which does not hold definitional evaluative features, gains axiological character, the result of which are extensions and semantic changes.

References

Bobbio, N. 1996. Left and Right. The Significance of a Political Distinction. Cambridge, Polity Press.

Brown, G. 1998. Context creation in discourse understanding. In Malmkjaer, K, Williams, J. (eds.) 1998, Context in Language Learning and Language Understanding, 171-193. Cambridge, Cambridge University Press.

Cruse, D. A. 2001. The Lexicon. In Aronoff, M., Rees-Miller J. (eds.) 2001, The Handbook of Linguistics. Oxford, Blackwell Publishing.

Dahl, R. 1995. Demokracja i jej krytycy. Kraków, Znak.

Yule, G. 1994. The study of language. Cambridge, Cambridge University Press.

Yule, G. 1996. Pragmatics. Oxford, Oxford University Press.

Language constructs as compatibility intervals: a small-scale experiment

Pavlo Kapustin[1], Michael Kapustin[2]
[1]Department of Information Science and Media Studies, University of Bergen, Norway
[2]Department of General Physics, Moscow Institute of Physics and Technology, Russia
https://doi.org/10.36505/ExLing-2020/11/0025/000440

Abstract

We conducted a small-scale experiment in which the respondents were asked to relate different language constructs (e.g. "possible", "afterwards", "seldom") to numbers on a certain scale (e.g. surprisingness, time, frequency). We studied how the respondents expressed vagueness of the meaning using a numeric scale, examined how their answers related to scalar implicatures and questioned whether the meaning of vague constructs could be adequately modelled using compatibility intervals, a meaning representation (computational model) we recently proposed.

Keywords: scales, vague constructs, meaning representations.

Introduction

Using numeric approaches for modelling meaning of vague constructs has been described, for example with use of relatively well-known fuzzy sets (Zadeh, 1971) or compatibility intervals (Kapustin and Kapustin, 2019a) that are closely related to fuzzy sets. Both meaning representations are based on the notion of compatibility of a given language construct (e.g. "young") with the values of a certain property (e.g. age) on a certain scale (e.g. 0-100).

While there are some examples in the literature that use these meaning representations, there is very little experimental work on the subject. Some studies show how such representations may be derived from experiments, e.g. (Hersh and Caramazza, 1976), however, to the best of our knowledge, no experiments aiming to directly express vagueness of meaning in a response using numbers have been performed.

We conducted a small-scale experiment exploring one possible approach to gathering such data. The experiment had the following goals:

- Study how the respondents would express vagueness of meaning using a numeric scale.
- Study how the respondents would compare meaning of related constructs using a numeric scale.
- Study whether the answers agree with scalar implicatures where appropriate.
- Study whether the meaning of vague constructs could be adequately modelled using compatibility intervals.

Compatibility intervals

Compatibility intervals consist of three number intervals: left ("increasing"), main and right ("decreasing"). For example, one could define the compatibility interval for "young" over property "age" as [0-18 – 30-50] (which is, of course, entirely subjective). This says that description "young" is compatible with ages 18-30, incompatible with ages over 50, and that the compatibility increases between ages 0 and 18 and decreases between ages 30 and 50. We use double hyphens between the start and the end of the main subinterval, and single hyphens between the start and the end of the left and the right subintervals.

Study design

We used an online survey to interview 22 respondents about 20 language constructs, asking the respondents to relate the constructs to 5 different properties (see Table 1). Possible answers were represented as numbered check boxes arranged horizontally to form a numeric scale. For each construct, the respondents were asked two questions, first focusing on the normal meaning of the construct and second focusing on a somewhat broader meaning of the construct (see an example question below). When it comes to the dimensions of meaning, we used one-dimensional projections, following some of the examples from Kapustin and Kapustin (2019b).

Properties and scales

Table 2. Properties and scales.

Property	Scale (possible answers)
Time	0 (right now) - 9 (in the distant future)
Surprisingness	0 (completely anticipated) - 9 (completely surprising)
Perceived duration	0 (instantaneous), 1(much shorter than expected), 2, 3, 4, 5 (as expected), 6, 7, 8, 9 (much longer than expected)
Event frequency	0 (never) - 9 (extremely often)
Event expectedness (probability in given conditions)	0 (not expected at all) - 9 (certain to happen)

Example question

Someone describes how often a certain event happens by using the following words. According to this description, how often can the event happen? Please select the choices that fit (0 means "never", 9 means "extremely often", and 1-8 correspond to everything in between).

a) According to the description "seldom", this event can occur: 0 (never), 1, 2, 3, 4, 5, 6, 7, 8, 9 (extremely often).
b) Somewhat less likely, the description "seldom" can also include events that occur: 0 (never), 1, 2, 3, 4, 5, 6, 7, 8, 9 (extremely often).

Results

We found that the majority of the responses formed coherent intervals, and most of them were also representable as compatibility intervals, e.g. see examples 1-3 in Table 2. While interesting, we think this is largely explained by the design of the experiment.

Some responses, however, could not be represented as compatibility intervals. In example 4 in Table 2, the answers to the first and second questions are far apart, possibly suggesting that the respondent has two different meanings of "entire" in mind. In example 5 in Table 2, the answer to the second question is a "shift" of the answer to the first question, possibly suggesting that the respondent gives an alternative, less "polar" meaning of "occasionally" as the answer to the second question. In total, there were very few answers similar to the last two examples. We believe this is partly because the phrasing of the questions implied that the answer to the second question should be a "wider" version of the same meaning of the construct.

Most of the respondents found that "seldom" is not compatible with events that never happen, always is not compatible with events that always happen, and "mere", "just" and "only" are not compatible with things that happen instantaneously (a somewhat invented example of such use would be "it only took me zero minutes to get there"). This agrees with scalar implicatures.

According to averaged responses, "later" [2.69-3.25–7.00-7.69] and "eventually" [3.76-5.12–8.35-8.53] have a wider spread than "afterwards" [1.13-1.60–4.13-5.00], and "eventually" spans further in the future than "later". Also, "regularly" seems to be a more concrete word in terms of event expectedness (probability in given conditions) than in terms of event frequency (averaged intervals [3.67-4.60–7.00-7.67] for frequency and [4.21-4.79–6.57-7.21] for expectedness). Also, "entire" [4.93-5.47–7.00-7.40] is a stronger construct than "whole" [4.80-5.07–6.33-6.80], when related to perceived duration.

Table 3. Answer examples.

	Construct	Property
1	**In a while** First question: 3, 4, 5, 6. Second question: 2, 7, 8 Compatibility interval: [2-3 – 6-8]	Time
2	**Possible** First question: 4, 5. Second question: 3, 4, 5, 6 Compatibility interval: [3-4 – 5-6]	Surprisingness
3	**Entire** (as in "an entire minute") First question: 7, 8, 9. Second question: 5, 6 Compatibility interval: [5-7 – 9]	Perceived duration
4	**Entire** (as in "an entire minute") First question: 5. Second question: 9	Perceived duration
5	**Occasionally** First question: 3, 4, 5. Second question: 4, 5, 6	Event frequency

Discussion

Despite the limitations of the current survey, we think that the results of the experiment are interesting. We find compatibility intervals to be a promising and relatively easily linguistically interpretable meaning representation that could be used for computational purposes. While specific compatibility intervals have no significance and are to a large degree context dependent, it is interesting how they compare to each other in individual responses, allowing to express certain aspects of the meaning of the language constructs in a quantitative, measurable way. We also think this perspective shows that such properties as vagueness and mirativity (Zeevat, 2013) may apply to more constructs than often considered, especially taking into account that many constructs have different dimensions of meaning, conveying information about different properties.

We think that further and improved experiments of this kind are interesting both from the computational and the theoretical perspectives. In the future, we would like to work on improving the survey, for example, add self-assessment of question understanding, randomize question order, use more constructs, more respondents and require the respondents to be native speakers of the survey language.

References

Hersh, H.M., Caramazza, A., 1976. A fuzzy set approach to modifiers and vagueness in natural language. Journal of Experimental Psychology: General 105, 254.

Kapustin, P., Kapustin, M., 2019a. Modeling language constructs with compatibility intervals, in: Proceedings of the IWCS 2019 Workshop on Computing Semantics with Types, Frames and Related Structures. pp. 49–54.

Kapustin, P., Kapustin, M., 2019b. Modeling language constructs with fuzzy sets: some approaches, examples and interpretations, in: Proceedings of the 13th International Conference on Computational Semantics-Student Papers. pp. 24–33.

Zadeh, L.A., 1971. Quantitative fuzzy semantics. Information sciences 3, 159–176.

Zeevat, H., 2013. Expressing surprise by particles, in: Beyond Expressives: Explorations in Use-Conditional Meaning. Brill, pp. 297–320.

Acoustic-orthographic interface in L2 phonology by L1 Cypriot-Greek speakers

Sviatlana Karpava[1], Elena Kkese[2]
[1]Department of English Studies, University of Cyprus
[2]Department of Education and Linguistics, UCLan, Cyprus
https://doi.org/10.36505/ExLing-2020/11/0026/000441

Abstract

The present study investigated the acoustic-orthographic interface in the phonology of L2 English by L1 Cypriot-Greek (CG) speakers. Seventy L1 CG undergraduate students completed a written dictation task, which examined how contrastive English vowels and consonants on word-level are perceived by CG and how the use of L2 affects these perceptions based on the different phoneme inventories and orthographies of CG and English. The findings suggest that there is an effect of L1 CG phonological and orthographic systems on L2 English vowel and consonant sound perception and written production.

Keywords: L2 phonology, perception, vowels, consonants, acoustic-orthographic interface

Effects of orthography on explicit phonemic processing

Orthographic forms or spellings are usually ignored in the L2 (second language) classroom context since the skill of spelling is believed to develop on its own (Kkese, 2020a). Nonetheless, L2 teachers have long known that orthographic forms can affect pronunciation since language learners are simultaneously exposed to the orthographic and phonological forms of the L2. Orthographic forms representing the sounds and/or words of a language in writing can affect language learners but also perception, production, and acquisition of L2 phonology and morphology.

With reference to the influence of orthography on L2 phonology, this can be positive facilitating L2 acquisition and pronunciation (Escudero et al., 2008); it can be negative leading to non-nativelike pronunciation (Bassetti, Atkinson, 2015; Young-Scholten, Langer, 2015); it can have mixed or no effects (Escudero, 2015). This happens because L2 learners have already acquired the phonological system and orthographic properties of the L1 (first language) and may draw on this knowledge while acquiring the target language (Kkese, 2020b).

L2 learners of English with L1 Standard Modern Greek (SMG) and/or Cypriot Greek (CG) background can have inappropriate inference from the orthography based on the wrong assumption that L2 English orthography is phonemic and there is a grapheme-phoneme correspondence (GPC) as in the

orthography of the L1 in which there is constant reading of graphemes (Kkese, 2020a).

The aim of this study is to investigate the acoustic-orthographic interface in L2 phonology by examining speech perception and written production of L2 English vowels and consonants by L1 CG users. As a result, for this study, the following research question will be investigated:

RQ1: Is there difference between perception and production of vowel and consonant phonemes in L2 English by L1 CG students? What is the effect of word frequency, the number of syllables in a word, the position of a sound in a word and the characteristics of acoustic input (male vs female voice)?

Methodology

Seventy CG undergraduate university students (1st year), who were learners of L2 English at a private, English-speaking university, participated in the study. There were 40 male and 30 female participants with normal speech and hearing. Their age ranged from 17 to 27 (Mean 19.8) and their L2 English proficiency was from low intermediate to advanced (5-9 IELTS scores, Mean 6.5).

In this study, a word dictation task was implemented to test L2 English vowel and consonant perception and production of sounds by L1 CG students, their word recognition skills, and accuracy of decoding acoustic speech signal into words. The dictation task had 120 test items: 60 for vowels and 60 for consonants. There were 10 conditions for consonant sounds (6 test items each): [ð], [z], [θ], [v], [d], [ŋ], [h], [b], [g], [ɪ] and 10 conditions for vowel sounds (6 test items each): [æ], [ɜ:], [ɔ:], [i:], [u:], [ɑ:], [e], [ʌ], [ə], [ʊ]. The dictation task was split into 6 dictation sessions; 20 test items for each (10 consonant and 10 vowel test items).

Results and Discussion

Overall, the results suggest that the dictation task seems to be quite difficult for students due to differences between the L1 CG and L2 English phonological and orthographical systems. There was a high percentage of no production and substitution errors. The participants had more non-target word transcription results for both vowels and consonants. This could also be due to a very strict scoring system: only accurate word recognition and word transcription was measured as target-like performance, see Table 1.

According to the paired samples t-test, there is a statistically significant difference between target and non-target production, for both consonants and vowels *(t(69)= -9.958, p=.000)*; vowel target and non-target production *(t(69)= -8.398, p=.000)*; consonant target and non-target production *(t(69)= -11.271, p=.000)*; and between target production of consonants and vowels *(t(69)= 6.492, p= .000)*.

Table 1. Vowel vs. consonant perception and production.

Overall	**Vowels**			**Consonants**		
target	48.67%			49.81%		
non-target	51.33%			50.19%		
Frequency	**high**		**Low**	**high**	**low**	
target	55.82%		41.51%	59.29%	40%	
non-target	44.18%		58.49%	40.70%	60%	
Position	**initial**	**middle**	**Final**	**initial**	**middle**	**final**
target	59.42%	44.51%	40%	54.31%	48%	45%
non-target	40.58%	55.50%	60%	45.69%	52%	55%
Syllables	**one**		**two**	**one**	**two**	
target	47.85%		46.89%	44.75%	52.90%	
non-target	52.15%		53.11%	55.25%	47.10%	
Voice	**Male**		**Female**	**male**	**female**	
target	45.16%		49.45%	50.46%	49%	
non-target	54.84%		50.56%	49.54%	51%	

Taking each condition separately into consideration, the participants had more target-like performance with respect to the following consonant sounds: [h] (85.18%), [ɹ] (70.20%), [z] (59.38%), [d] (54.44%) and [b] (56.74%) while the most vulnerable conditions were with the consonant sounds: [ð] (23.13%), [θ] (37%), [v] (42.91%), [ŋ] (29%) and [g] (40.07%).

These findings are in agreement with Kkese, Karpava (2019), though the experimental tasks were different. The participants had mainly substitution and no production errors regarding consonant conditions. The substitution errors were based on the similarity or contrast of voice feature and manner of articulation. There is a strong effect of word frequency on target perception and production of consonant sounds as high-frequency words elicit more target-like answers. Overall, high-frequency words have an advantage over low-frequency words in terms of the number of acoustic cues accumulated over time.

The students were better in terms of comprehension and production of the following vowel sounds: [ə] (schwa) (74.62%), [uː] (64.19%), [ɔː] (53.81%) and [e] (52.86%) than the rest of the vowel sounds: [æ] (44.17%), [iː] (47.68%), [ɑː] (42.47%), [ʌ] (40.62%), [ʊ] (44.35%). The most vulnerable condition was [ɜː] (21.90%). The high production of the schwa sound [ə] can be explained by the fact that it is the most common vowel sound in English, which can be spelled with any vowel grapheme. It appears on unstressed syllables and has a neutral mouth position.

This can be due to the differences between English and CG phonological and writing systems; in CG, there is no long-short vowel distinction while the

schwa sound is absent. Non-target perception and production of vowels was characterised by no production and substitution errors based on the similarity of [±back], [±round] and duration features. There was a clear effect of the word frequency on vowel perception and production as high-frequency words triggered more target-like test performance. Other factors, such as number of syllables in a word, position of the sound in a word, and acoustic input characteristics do not influence the written transcription of L2 English words.

Both consonant and vowel data of this study support the idea of Bassetti, Atkinson (2015) and Young-Scholten and Langer (2015) about the negative inter-orthographic effects on L2 phonological representations and L2 sound perception and production.

References

Bassetti, B., Atkinson, N. 2015. Effects of orthographic forms on pronunciation in experienced instructed second language learners. Applied Psycholinguistics 36, 67-91.

Escudero, P. 2015. Orthography plays a limited role when learning the phonological forms of new words: The case of Spanish and English learners of novel Dutch words. Applied Psycholinguistics 36, 7-22.

Escudero, P., Hayes-Harb, R., Mitterer, H. 2008. Novel second-language words and asymmetric lexical access. Journal of Phonetics 36, 345-360.

Kkese, E. 2020a. L2 writing Assessment: The Neglected Skill of Spelling. Newcastle upon Tyne, Cambridge Scholars Publishing.

Kkese, E. 2020b. Phonological awareness and literacy in L2: Sensitivity to phonological awareness and phoneme-grapheme correspondences in L2 English. In Neokleous, G., Krulatz, A., Farrelly, R. (eds.) 2020, Handbook of Research on Cultivating Literacy in Diverse and Multilingual Classrooms. Hershey, PA, USA, IGI Global Press.

Kkese, E., Karpava, K. 2019. Applying the Native Language Magnet Theory to an L2 setting: Insights into the Cypriot Greek adult perception of L2 English. In E. Babatsouli, E. (ed.), Proc. of the International Symposium on Monolingual and Bilingual Speech 2019, 67-74, Chania, Greece.

Young-Scholten, M., Langer, M. 2015. The role of orthographic input in second language German: Evidence from naturalistic adult learners' production. Applied Psycholinguistics 36, 93-114.

Disambiguation in corpus of Modern Greek

Maxim Kisilier[1,2], Olga Nikolaenkova[1]
[1]Department of General Linguistics, Saint Petersburg University, Russia
[2]Institute for Linguistic Studies, Saint Petersburg, Russia
https://doi.org/10.36505/ExLing-2020/11/0027/000442

Abstract

Corpus of Modern Greek appeared in 2011. All texts are morphologically annotated. Due to certain peculiarities of Modern Greek morphology, the majority of forms has more than one grammatic interpretation. In this presentation we describe the types of homonyms which are found in the Corpus and discuss possible patterns for automatic disambiguation. At the end, we mention a number of problematic cases that cannot be resolved now or require manual approach.

Keywords: language corpus, ambiguity, Modern Greek, automatic disambiguation

General remarks

Corpus of Modern Greek (= CMG, http://web-corpora.net/GreekCorpus/search/?interface_language=en, access date 31.08.2020) was created in 2011 with the support of the "Corpus linguistics" program of the Russian Academy of Sciences. The size of CMG is over than 35.5 million tokens and some of its functions are absent from other corpora of Modern Greek (cf. Arkhangelskiy & Kisilier 2018). All texts in CMG are morphologically annotated by means of a digital grammatical dictionary and morphological analyzer (UniParser). The set of morphosyntactic values used for annotation coincide with basic grammatical categories (gender, number, case, tense, etc.).

Since the morphological annotation is an automatized process, each word has all possible analyses. Unlike Ancient Greek, flexions in Modern Greek often do not provide enough information to distinguish different forms (for example, αδελφ-ή 'sister' may be both nominative and accusative), and thus the percentage of ambiguous words and forms is high — according to Elizaveta Kuzmenko & Elmira Mustakimova (2015: 390) it is approximately 43%. It may grow with the further development of CMG and it is important to elaborate the mechanisms of disambiguation. Even now (when the corpus is not so large) manual disambiguation is not possible and at least the most typical cases should be disambiguated automatically.

Previous attempt of automatic disambiguation (Kuzmenko & Mustakimova 2015: 390) took into account only some definite articles, personal pronouns and certain forms of the verb 'to be'. Formally, these are the most frequent ambiguity examples in CMG, but from the point of view of Modern Greek morphology, they are less systematic than the homonymy of morphological

flexions. We believe that disambiguation in CMG requires a more systematic approach and, in this presentation, we intend to describe a number of most typical ambiguities and to discuss which of them do not require manual work. Most examples used in this paper are from CMG.

Lexical ambiguity

Lexical ambiguity, or homonyms is the best-known type of ambiguity and it is widespread in Modern Greek: βήμα — (a) 'step, pace', (b) 'tribune'. Such examples do not require any disambiguation at all as they do not affect morphological annotation.

Semilexical ambiguity

Semilexical homonyms usually belong to different morphological categories or classes:

(1) του (a) article in genitive
(b) personal clitic pronoun in genitive
(c) possessive pronoun

Their disambiguation could be based on syntax restrictions: (a) article always precedes NP, while (b) personal clitic pronoun is in front of a finite verb or after imperative/participle and (c) possessive pronoun follows either a noun or an adjective.

Sometimes situation, at first sight, looks more complicated:

(2) η **θεία** Αγάπη
ARTICLE aunt.NOUN Agapi
ARTICLE divine.ADJECTIVE.FEMININE love
'**aunt** Agapi (personal name)' or '**Divine** Love'

So far, (2) has no solution. Even a not fully proficient speaker of Modern Greek may get confused here. But let us take a look at (3) and (4) which illustrate the most typical usage of these homonyms:

(3) η **θεία** Ιουλία
ARTICLE aunt.NOUN Julia
'**aunt** Julia'

(4) η **θεία** λειτουργία
ARTICLE divine.ADJECTIVE.FEMININE liturgy
'**divine** liturgy'

Evidently, the noun (θεία/θείος 'aunt/uncle') is more commonly used both with a personal name or independently and the adjective is likely to be accompanied with a common name. If a number of semantic values is added to the grammatical dictionary in CMG, automatic disambiguation will be based on syntactic/combinatory restrictions. Certainly, some problematic situations, like (2), will not be resolved but their number will hardly exceed 2 or 3%.

Morphological ambiguity

This class includes several declension types where some flexions of different cases coincide, for example:

(5)

	SINGULAR		PLURAL
	NOMINATIVE	μητέρ-**α** 'mother'	μητέρ-**ες**
	ACCUSATIVE	μητέρ-**α**	μητέρ-**ες**

(6)

	SINGULAR		PLURAL
	NOMINATIVE		ψαράδ-**ες**
	GENITIVE	ψαρ-**ά** 'fishmen'	
	ACCUSATIVE	ψαρ-**ά**	ψαράδ-**ες**

Although (5) and (6) represent different declensions, disambiguation mechanism for all feminine and masculine nouns will be the same — the article will always indicate the right case. It is important to take into account that the article may be placed distantly:

(7)

η	**όμορφη**	γυναίκα	
ARTICLE	beautiful	woman	
'beautiful woman'			

(8)

η	**αγαπημένη**	**μου**	γυναίκα
ARTICLE	beloved	my	woman
'my beloved woman'			

(9)

η	**πολύ**	**αγαπημένη**	**μου**	γυναίκα
ARTICLE	very	beloved	my	woman
'my very much beloved woman'				

It is not very difficult to define the list of constituents which may separate the article from the noun (adjective/participle, possessive pronoun, few adverbs, etc.) even despite the fact that some types of constituents may be used more than once:

(10)

η	πολύ	**αγαπημένη**	μου	και	**ξεχωριστή**	γυναίκα
ARTICLE	very	beloved	my	and	exceptional	woman
'my very much beloved and special woman'						

Problems

One of the major challenges we face with neuter where the article does not help to distinguish nominative from accusative:

(11)

	SINGULAR	PLURAL
	NOMINATIVE	**το** λουλούδ-**ι** 'flower'	**τα** λουλούδ-**ια**
	ACCUSATIVE	**το** λουλούδ-**ι**	**τα** λουλούδ-**ια**

Modern Greek is a free word order language, that is why a syntactic regulation is not applicable here.

Another difficulty for automatic disambiguation are conjunctions ότι, που and πως which may be either complementizers or not. In (12), ότι is not a complementizer ('that') but an anaphoric pronoun (= ο τι):

(12) λέγε **ὁτι** θες
say what you.want
'say **whatever** you want'

However, only intonation or wider context helps to understand it. The same is relevant for another conjunction που:

(13) σε εσένα το λέω **που** μαζί τους συμφωνείς
to you it I.say who/that with them you.agree
'I say it to you **who** agree with them'

In (13), που does not refer to the adjacent verb but to pronoun εσένα.

(14) έλεγα <…> **πως** πολύ μου αρέσει,
I.said that/how much I like
πως είμαι περίεργος
that/how I.am curious

Both πως in (14) depend on the verb έλεγα despite the fact that the second πως immediately follows the verb αρέσει. Still without a wider context or intonation it is not clear whether πως means 'that' ('I said <…> **that** I like [it] very much, **that** I am curious [about it]') or 'how' ('I said <…> **how** much I like [it], **how** curious I am').

Certainly, there are some limitations in the use of complementizers, but the recent corpus-based analysis in (Kisilier 2020) clearly demonstrates that the system of complementizers is rapidly changing. Probably the best solution is to accept that in Modern Greek ὁτι, που and πως have multiple coexisting meanings which refer to the same word and are not homonymic.

Acknowledgements

This research was supported by the Russian Foundation for Basic Research (Project No 18-012-00607 "Corpus and fieldwork-based studies of sentential complements in the Balkan languages and dialects").

References

Arkhangelskiy, Timofey & Maxim Kisilier. 2018. Corpora of Modern Greek: achievements and goals (Корпуса греческого языка: достижения, цели и задачи). Indo-European Linguistics and Classical Philology 22(1), 50–59.

Kisilier, Maxim. 2020. Ὁτι, που and πως in Standard Modern Greek (Комплементайзеры ὁτι, που и πως в новогреческом языке). Indo-European Linguistics and Classical Philology 24(1). 554–577.

Kuzmenko, Elizaveta & Elmira Mustakimova. 2015. Automatic disambiguation in the corpora of Greek and Yiddish. In Computational linguistics and intellectual technologies (Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (2015) 1.) Annual International Conference "Dialogue". Vol. 1, 388–397. Moscow: Russian State University of Humanities.

Thematic role and grammatical function affect pronoun production

Sofiana-Iulia Lindemann, Stanca Mada, Laura Sasu, Madalina Matei
Department of Theoretical and Applied Linguistics, Transilvania University of Brasov, Romania
https://doi.org/10.36505/ExLing-2020/11/0028/000443

Abstract

According to different approaches to pronoun processing, in pro-drop languages, null pronouns are interpreted as referring back to the grammatical subject and topical referent, while overt pronouns are usually interpreted as coreferring with a non-subject and non-topical antecedent. The present study investigates whether thematic role and grammatical function impact (overt and null) pronoun production in Romania. Results show that we do not encounter a clear division of labour between the two pronoun forms triggered by syntactic structure alone and that thematic roles matter as well. The findings support a multi-dimensional approach, suggesting that different referential forms are constrained by different factors.

Keywords: Thematic role, grammatical function, pronoun interpretation, Romanian

Introduction

Romanian is a pro-drop language (Dobrovie-Sorin 1994), which allows for both overt and covert pronouns in preverbal subject position. According to different studies, null pronouns are used for the most prominent antecedent, while overt pronouns are employed to refer back to a less prominent antecedent (Ariel 1990, Carminati 2002). The prominence or accessibility status of a given referent is affected by grammatical function. Accordingly, the grammatical subject referent is considered prominent and is realized by means of a null pronoun. Other referents that are realized in less prominent positions (e.g. objects) are associated with more elaborated types of referring expressions, such as overt pronouns, definite noun phrases and names.

Besides grammatical function, other factors have been shown to impact the prominence status of a referent and, consequently, the type of referring expression used to refer back to that referent, such as recency, parallelism, topicality (Ariel 1990, Arnold 2001, Chiriacescu 2011b).

Most studies investigating the distribution of pronouns in Romanian were comprehension studies focusing on intra-sentential contexts, exploring whether a null or overt pronoun co-refers with a previously introduced subject or object referent. In this study, we are interested in production preferences: when do writers use an overt pronoun over a null one? And, do the observed biases towards using a pronoun form hold across sentence boundaries? Furthermore,

the present paper focuses on thematic roles and whether they have an impact on pronominalization as well, besides the attested grammatical role preference. In the next sections we present the sentence completion study followed by the conclusions.

The experimental study

The aim of the present study is to extend the existing literature on accessibility and prominence and, more importantly, to investigate whether thematic role and grammatical function impact (overt and null) pronoun production in Romania.

Participants

110 native Romanian speakers from the Transilvania University of Brasov, Romania, took part in the experimental study (age range: 18-45 years, mean age 28 years, 65 female). It took about twenty minutes to complete each version of the study.

Design, procedure and materials

A story completion task was used in which each target sentence consisted of two human referents that had the same gender. As speakers prefer pronoun anaphora when referring back to a previously introduced subject, we used transfer-of-possession-verbs as they allow for the separation of grammatical roles from thematic roles. Some of these verbs realize the Goal as the grammatical subject (e.g. *get*, CND_3 and CND_4 in Table 1), while others realize the Source in grammatical subject position (e.g. *give*, CND_1 and CND_2 in Table 1). Participants' task consisted in reading the given one-sentence target items and providing natural sounding written continuations to each sentence they read. Furthermore, we specifically asked participants to re-mention the underlined referent in their continuation sentence (method adapted from Fukumura & van Gompel 2010). This manipulation resulted in 4 conditions, crossing thematic roles (Goal vs. Source) and grammatical function (Subject vs. Object). We used 36 experimental items, 9 for each condition, and 40 filler items, distributed in two lists.

Table 1. Conditions from the experimental study with English translations.

CND1_Goal=Object	Paul i-a dat un bilet <u>lui Andrei.</u> Paul gave a note <u>to Andrew.</u>
CND2_Source=Subject	<u>Paul</u> i-a dat un bilet lui Andrei. <u>Paul</u> gave a note to Andrew.
CND3_Goal=Subject	<u>Raul</u> a cumpărat o carte de la Flavius. <u>Raul</u> bought a book from Flavius.
CND4_Source=Object	Raul a cumpărat o carte <u>de la Flavius.</u> Raul bought a book <u>from Flavius.</u>

Results

The main aim of the study was to assess whether thematic roles affect the production of different types of referring expressions in Romanian alongside the attested grammatical bias. Two independent judges coded for the type of referring expression chosen by the participants to refer back to the underlined referent (i.e. proper name, definite noun phrase, overt pronoun, null pronoun, other). We coded 900 continuations.

Overall, participants used more null pronouns to refer back to the subject referent rather than the non-subject. Overt pronouns were used for both subject and non-subject referents. The preference of the null pronoun to pick up the subject is stronger than the preference of the overt personal pronoun to pick one of the two referents. The findings indicate that we do not find a clear division of labour between the two pronoun forms in terms of syntactic structure.

Interestingly, we found an additional effect of thematic role. This effect was particularly evident for null pronouns, which were more likely to pick up Goal referents than Source referents in subject position. Thus, we found the highest rates of null pronoun use for CND3_Goal=Subject followed by CND2_Source=Subject. In this study, overt pronouns were used more often for the second most prominent referent in terms of both syntactic structure and thematic role. Overt pronouns were used most often for the underlined referents in the CND2_Source=Subject condition, followed by referents introduced in the CND1_Goal=Object condition. Overall, we found most proper names used to pick up the referent in CND2, which was realised as the thematic Source and as the grammatical subject.

Conclusions

The experimental study revealed two main findings. First, grammatical role has a strong impact on the type of referring expression used. This is in line with previous results that showed that speakers tend to avoid using more elaborated types of referring expressions like names or definite noun phrases for referents recently mentioned in subject position (Ariel 1990, Arnold 1998, Gundel et al. 1993, Chiriacescu 2011b). Instead, they choose reduced types of referring expressions like (overt or null) pronouns.

Second, results showed that the use of a particular thematic role does matter, as it affects the choice of subsequent mention as well (Rosa & Arnold 2017). Being realised as the grammatical subject referent and simultaneously as the thematic Goal increases the rate of subsequent null pronoun use. The same observation holds for non-subjects realized as the thematic Goal, for which more overt pronouns are subsequently used.

In sum, the present results support a multi-dimensional approach, suggesting that different referential forms are constrained by different grammatical and

semantic factors (Kaiser & Trueswell 2008; con Heusinger & Chiriacescu 2009, Chiriacescu 2011a).

Acknowledgements

This research was supported by a grant of the Romanian Ministery of Research and Innovation, CNCS – UEFISCDI to project PN-III-P1-1.1-TE-2016-1241, within PNCDI III.

References

Ariel, M. 1990. Accessing noun-phrase antecedents. London, Routledge.

Arnold, J.E. 2001. Reference form and discourse patterns. Unpublished doctoral dissertation. Stanford University, Stanford, CA.

Carminati, M.N. 2002. The processing of Italian subject pronouns. Doctoral dissertation. Retrieved from Proquest (AAI3039345).

Chiriacescu, S.I. 2011a. Effects of reference form on frequency of mention and rate of subsequent mention, In S.L. Devi, A. Branco and R. Mitkov (eds.) 2011, Anaphora Processing and Applications. Lecture Notes in Computer Science, 132-142. Heidelberg, Springer.

Chiriacescu, S.I. 2011b. Factors contributing to the salience of referents. In M. Nevaci (ed.) 2011, Studia linguistica et philologica, 74-96. Bucuresti, Editura Universitatii.

Dobrovie-Sorin, C. 1994. The Syntax of Romanian. Berlin, De Gruyter.

Fukumura, K, van Gompel, R.P.G. 2010. Choosing anaphoric expression. Do people take into account likelihood of reference? Journal of Memory and Language 62, 52-66.

Gundel, J., Hedberg, N, Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. Language 69(2), 274–307.

von Heusinger, K., Chiriacescu, S.I. 2009. Definite "bare" nouns and pe-marking in Romanian. In Espinal, M.T., Leonetti, M., McNally, L. (eds.) 2009, Proceedings of the IV Nereus International Workshop "Definiteness and DP Structure in Romance Languages", 63-82. Konstanz, Universität Konstanz, Arbeitspapier 124. Fachbereich Sprachwissenschaft.

Kaiser E., Trueswell, J. 2008. Interpreting pronouns and demonstrantives in Finnish: Evidence from a form-specific approach to reference resolution. Language and Cognitive Processes 23(5), 709-748.

Rosa, E., Arnold, J. 2017. Predictability affects production: Thematic roles can affect reference form selection. Journal of Memory and Language 94, 43-60.

Acoustic characteristics of Italian Parkinsonian speech: a study on early-stage patients

Marta Maffia[1], Rosa De Micco[2], Alessandro Tessitore[2], Anna De Meo[1]

[1]Department of Literary, Linguistics and Comparative Studies, University of Naples "L'Orientale", Italy

[2]Department of Advanced Medical and Surgical Sciences, University of Campania "Luigi Vanvitelli", Italy

https://doi.org/10.36505/ExLing-2020/11/0029/000444

Abstract

Parkinson's Disease (PD) is a neurological illness which also has effects on speech production, resulting in segmental and suprasegmental abnormalities. The aim of the current study is to test the validity of two acoustic parameters - %V and VtoV - for the detection of rhythmical variation in early-stage PD speech, in comparison to healthy speech. 40 Italian native speakers were enrolled in the research, 20 early-stage PD subjects and 20 neurologically healthy and matched controls, and a corpus of read speech was collected. The results of voice analysis confirmed an alteration of vocalic duration and %V in PD productions. In particular, %V could be a reliable cue for PD speech characterization, even at the very early onset of the disease.

Keywords: Parkinson's disease, language pathology, acoustic analysis, speech rhythm, early diagnosis

Introduction

Parkinson's Disease (PD) is recognized as the second most common neurodegenerative disorder after Alzheimer's Disease, with a point prevalence ranging from 0.25-4% between age of 65-80 (de Lau & Breteler 2006).

Degeneration of nigrostriatal dopaminergic neurons, which results in disruption of basal ganglia–thalamo-cortical loops, underlies the classical motor signs and symptoms of PD (i.e. bradykinesia, rigidity, tremor and postural instability). As reported in experimental studies on different languages, PD-related dysarthria also has effects on the three major anatomic subsystems governing speech motor control: respiration, phonation and articulation. At both the segmental and the suprasegmental level of speech, the hypokinesia and the resulting decreased amplitude of motion of lips, tongue and jaw may cause a variety of deficits, such as imprecise vowel and consonant articulation, reduced or abnormal pitch variability, speech fluency and speech rate (see, among the others, Skodda et al. 2012, Liss et al. 2009).

In this framework, it has been suggested that the observation of some acoustic parameters' variation may potentially provide a sustainable and non-invasive diagnostic tool, in support of clinical assessment, even in the very early stages of the disease, when the neurodegeneration is yet started and spread throughout the nervous system but still there are no motor signs (Harel et al. 2004).

Previous researches on different languages and on mild-to-severe PD patients have shown the effectiveness of two acoustic parameters, vowel percentage (%V) and the mean interval between two consecutive vowel onset points (VtoV), in the description of rhythmic variations of PD speech, compared to healthy speech (Pettorino et al. 2016, Pettorino et al. 2017).

The purpose of the present study is to verify the validity of the two acoustic parameters, %V and VtoV, even for the speech characterization of Italian subjects with early-stage PD.

Methods

Subjects

The data for the present research were collected from 40 speakers: 20 individuals diagnosed with PD and 20 age and sex-matched healthy controls (HC) with no history of neurological or speech disorders. All subjects were Italian native speakers residing in Campania region and they all gave written consent to the data collection procedure.

Patients were recruited from an ongoing longitudinal study enrolling non-demented early PD patients diagnosed according to the modified diagnostic criteria of the UK Parkinson's Disease Society Brain Bank (Gibb & Lees, 1988) at the Movement Disorders Unit of the First Division of Neurology at the University of Campania "Luigi Vanvitelli" (Naples, Italy). They underwent an extensive motor and non-motor assessment with validated PD-related scales.

Speech data collection and analysis

Each subject was instructed to read aloud from a printed sheet an expository text (about 350 syllables), chosen for having a low level of syntactic and lexical complexity and a very simple content. They were encouraged to speak in their normal, conversational voice, as spontaneously as possible and at comfortable loudness. The speech samples were recorded on a standard personal computer in a quiet room of University "Luigi Vanvitelli", by means of the software Praat (Boersma & Weenink 2009) at a 44100 Hz sampling rate. Sociolinguistic information on each speaker were also obtained with a questionnaire.

Patients performed speech assessment while taking their regular dopaminergic medications.

All the read speech samples were manually labelled to identify consonantal and vocalic intervals, through the visual inspection of speech spectrograms and waveforms. Approximants were completely avoided in the chosen text, while diphthongs were considered as a single vocalic interval. In V+nasal consonant

sequences, the nasalized portion of the vowel was assigned to the V interval. As for initial voiced stop consonants, the first boundary was considered to be the onset of the glottal pulses. Post-pausal voiceless plosives were assigned a duration equal to the mean value of single plosives in the same utterance.

Once extracted the durations of all consonantal and vocalic portions, %V and VtoV were obtained using a Praat script. Disfluences and silent pauses were not considered for the calculation of these two parameters.

Statistical analysis of clinical and speech data

T-test or X^2 were computed to compare clinical and speech data as appropriate. A $p<0.05$ was considered as statistically significant. Analyses were performed with SPSS version 23.

Results

The results of acoustic analysis are reported in Table 1, along with demographic and clinical data for all participants.

Table 1. Demographic, clinical and acoustic features of HC and PD patients.

		HC (n=20) mean±SD	PD (n=20) mean±SD	p-value
Demographic data	Age	64.8±5.9	63.8±10.9	0.713
	Sex (M/F)	8/12	12/8	0.205
	Disease duration (months)	-	31.7±18.9	-
Clinical data	H&Y stage	-	1.9±0.5	-
	UPDRS III	-	25.0±8.2	-
	MoCA total	-	24.4±4.3	-
	BDI	-	5.2±6.2	-
	LEDD total	-	289.9±152.3	-
	LEDD DA	-	78.3±129.1	-
Acoustic data	durV (ms)	81.6±7.9	91.3±12.1	**0.004**
	durC (ms)	103.2±9.1	93.5±12.7	**0.007**
	%V	44.7±1.3	49.8±1.7	**<0.001**
	VtoV	0.186 ±0.02	0.185±0.02	0.889

H&Y: Hoehn & Yahr; UPDRS: Unified Parkinson's Disease Rating Scale; MoCA: Montreal Cognitive Assessment; BDI: Beck depression inventory; LEDD: Levodopa Equivalent Daily Dose; DA: dopamine-agonist. Significant differences are reported in bold.

The %V is, thus, confirmed to be the parameter that changes most significantly in the two groups of speakers, being distinctly higher in PD speech (from 47.8% to 54.2%) than in HC productions (from 42.3% to 47.2%). Despite a high variability across speakers within both groups, on average, consonants are shortened in PD speech, as a possible consequence of consonant weakening, while vowels are significantly longer in PD productions than in HC speech.

With regard to VtoV, the two groups of participants present very similar mean values and the same speech rate variability.

Discussion and conclusions

First results suggest that %V is specifically altered in PD patients even at the disease onset. Dopaminergic pathways, which are known to be primarily affected at PD onset, may be also crucially involved in speech processes, altering very early the articulatory passage from the static phase, represented by vowels, to the dynamic phase of consonant production. This would account for the rhythmic characterization of PD speech, in which more time is spent on vocalic sounds with respect to healthy controls. Further investigations on larger PD samples are needed to support our observations.

Acknowledgements

We would like to thank Prof. Gioacchino Tedeschi and Prof. Massimo Pettorino for their precious support.

References

Boersma, P., Weenink, D. 2020. Praat: doing phonetics by computer [Computer program]. Version 6.1.20beta, retrieved 10 September 2020 from http://www.praat.org/.

de Lau, L. M., Breteler, M.M. 2006. Epidemiology of Parkinson's disease. Lancet Neurol 5, 525–535.

Gibb, W.R., Lees, A.J. 1988. A comparison of clinical and pathological features of young- and old-onset Parkinson's disease. Neurology 38, 1402–1406.

Harel, B.T., Cannizzaro, M.S., Cohen, H., Reilly, N., Snyder, P.J. 2004. Acoustic characteristics of Parkinsonian speech: a potential biomarker of early disease progression and treatment. Journal of Neurolinguistics 17, 439–453.

Liss, J.M., White, L., Mattys, S.L., Lansford, K., Lotto, A.J., Spitzer, S.M., Caviness, J.N. 2009. Quantifying speech rhythm abnormalities in the dysarthrias. Journal of Speech, Language, and Hearing Research 52, 1334–1352.

Pettorino, M., Busà, M.G., Pellegrino, E. 2016. Speech Rhythm in Parkinson's Disease: A Study on Italian. Proc. of the 17th Annual Conference of the International Speech Communication Association, San Francisco, USA, 1958-1961.

Pettorino, M., Gu, W., Pólrola, P., Fan, P. 2017. Rhythmic Characteristics of Parkinsonian Speech: A Study on Mandarin and Polish. Proc. of the 18th Annual Conference of the International Speech Communication Association, Stockholm, Sweden, 3172-3176.

Skodda, S., Gronheit, W., Schlegel, U. 2012. Impairment of vowel articulation as a possible marker of disease progression in Parkinson's Disease, PLoS ONE 7/2, 1-8.

Automatic detection of accent phrases in French

Philippe Martin
LLF, UFRL, Université de Paris
https://doi.org/10.36505/ExLing-2020/11/0030/000445

Abstract

In lexically-stressed languages such as English or Greek, accent phrases usually include one lexical word (noun, verb, adverb or adjective), together with some syntactically bound grammatical words (conjunction, pronoun or preposition). In non-lexically languages such as French or Korean, accent phrases are delimited by a final syllabic stress and may contain more than one lexical word, depending on the speech rate and limited to a 250 ms to 1250-1350 ms duration range. As perception of syllabic stress is strongly influenced by the listeners current own speech rate making perception agreement between annotators elusive, an interactive software program has been implemented imbedding constrains external to acoustic data to better investigate the actual distribution of stressed syllables in oral recordings of French.

Keywords: accent phrase, French, speech rate, WinPitch

Introduction

When we speak or read a text, we process speech not word by word, but accent phrase by accent phrase. Each accent phrase may include a single word, a group of words, or even a single syllable, but does contain only one stressed syllable (excluding emphatic stress). In languages such as English, Italian or Greek, this unique stressed syllable is located on a so-called lexical or content word, i.e. a noun, a verb, an adverb or an adjective, and its position is defined by the lexicon. Pronouns, conjunctions and articles belong to the category of grammatical words and are not normally stressed. An accent phrase therefore consists of one content word together with syntactically bound grammatical words.

The segmentation into accent phrases is the first step to recover the prosodic structure intended by the speaker (or the writer) in order to bootstraps the discovery of the syntactic structure. Even when we read a text silently ("in our head"), we do generate mentally a prosodic structure organizing the accent phrases in a hierarchy to access the meaning of the text, although in this case any acoustic input is of course absent.

French, however, does not have lexical stress but a group stress located on the last syllable of accent phrases. Rather than systematically constitute *groupes de sens*, accent phrases ended by syllabic stress are constrained by external rhythmic constrains, among which:

a. The minimal duration between consecutive stressed syllables, about 250 ms (Martin 2014).
b. The maximum duration between consecutive stressed syllables in continuous speech, about 1250-1350 ms (Martin 2014).
c. The eurhythmic tendence to balance the duration of successive accent phrases (Wioland 1983).
d. The perception of any word final syllable followed by at least 250 ms of silence as stressed (Martin 2018).

The only acoustic condition for a syllable to be identified as stressed pertains to the change of its vowel fundamental frequency perceived as a change in frequency rather than a static tone, i.e. above the glissando threshold (Rossi 1971).

The simple fact that we can restore stress locations when we read aloud or silently suggests that we may not really need any acoustical input to perceive stressed syllables when we listen to speech (again non-emphatic). The predictability of lexical stress leads to consider the perception of stressed syllables as a process which compares the actual acoustic features of syllables with a position predicted from the listener knowledge of accent phrases. Furthermore, eurhythmicity makes the perception of stressed syllables and the segmentation into accent phrases sensitive to listeners speaking or reading rate.

Automatic annotation of stressed syllables

The annotation of stressed syllable in speech corpora analysis is an essential step especially for macrosyntactic description of spontaneous speech based on accent phrase chunks. The perception of stress will be influenced by the annotator own prediction process, and stressed syllables located where they would have placed by reading or speaking at the annotator own pace. The problem for an annotator is therefore to adapt to the speech rate of the recording when stressed syllables are annotated (Martin 2020).

Automatic detection of stressed syllables in French operates usually in a bottom-up fashion from the speech acoustic data, looking for significant variations between consecutive syllables in duration, fundamental frequency and intensity. Vowel quality does not appear as a significant parameter for stress detection in French.

With the advent of relatively large corpora of both read and spontaneous speech, more systematic experimental studies were carried out, only the highlight the confusion in the domain. A paper from Avanzi (2013) for instance, faced with the uncertainty in annotating stressed syllables in French, describes in detail a complex procedure involving two experts. Even with this protocol, agreement between annotators varies between 60 % and 80 %.

Later, Christodoulides and Avanzi (2014) implemented an automatic detector of prominence (i.e. not just accent phrases stressed syllable) by

machine learning methods applied to a large corpus which included two different styles. Although they use a comprehensive set of acoustic parameters, their best results, evaluated against manual placement by experts in syllabic prominence, reaches a 90% correct identification level.

Considering these difficulties, it appears that stress detection should proceed not only from speech wave acoustic analysis, but also from the implementation of the rhythmic constrains evoked above.

Interactive computer annotation of stressed syllables

Rather than proceed by annotating stressed syllables by experts, or develop another algorithm proceeding from the speech signal, an interactive software program has been implemented in the WinPitch package. The idea was to automatically locate stressed syllables from the rules given above, each rule being adjustable by the user in order to evaluate their effect. It follows that missing or superfluous cases are quickly detected visually, thanks to a specific color code.

The implemented analysis proceeds as follows:

- Loading of a recording to analyze, with orthographic transcribed sections (each section of about 1 to 5 seconds).
- Automatic segmentation into vowels and consonants based on forced alignment with a text-to-speech segment generated from the annotated text.
- Automatic scanning and detection of stressed syllables (actually stressed vowels) based on the criteria enounced above:
  a. Any syllable followed by more than 250 ms silence is stressed
  b. Any final syllable of a noun, adjective, verb or adverb is stressable (from original accent phrase definition)
  c. If 2 consecutive stressed syllables are separated by less than 250 ms, the first one is unstressed (accent phrase minimum duration)
  d. Any stressable syllable with F0 change over the glissando threshold is stressed
  e. If 2 consecutive stressed syllables are separated by more than 1250 ms in continuous speech, at least one stressable syllable in this interval is stressed (accent phrase maximum duration). Make stressed the one with the highest glissando value
  f. One stressable syllable must exist in any time window duration equals to the accent phrase average duration (eurhythmy)

All parameters, glissando threshold, minimum and maximum accent phrase syllable… are user adjustable. The eurhythmic aspect is implemented by evaluating the first accent phrases realizations and the number of syllables they contain. This starting accent phrase duration will them be used to define a

sliding time window, in which most prominent syllables with a glissando value above the threshold are retained as stressed. The size of this sliding window defines a speech rate assumed to be constant for the whole recording.

Conclusions

Rather than being simply another stressed syllable detection algorithm, the automatic accent phrase detection process implemented in the speech analysis software WinPitch appears as an exploration tool allowing to quickly and efficiently test the rhythmic and acoustic parameters of stress in French. Both large read and spontaneous corpora have been processed, establishing the validity of these parameters, and more important, highlighting a perception process of accent phrases stressed syllables operating in both bottom-up and top-down mode.

References

Avanzi, M. 2013. « Note de recherche sur l'accentuation et le phrasé à la lumière des corpus du français », Tranel, vol. 58, 5-24.

Christodoulides, G., Avanzi, M. 2014. An Evaluation of Machine Learning Methods for Prominence Detection in French, Proc. Interspeech 2014, 116-119.

Martin, Ph. 2014. Spontaneous speech corpus data validates prosodic constraints, Proc. Speech Prosody 2014, 525-529.

Martin, Ph. 2018. Intonation, structure prosodique et ondes cérébrales, London: ISTE, 322 p.

Martin, Ph. 2020. L'annotation prosodique dans Orfeo, Langages 2020/3 (N° 219), 103-115.

Rossi, M. 1971. Le seuil de glissando ou seuil de perception des variations tonales pour la parole. Phonetica. n° 23, 1-33.

Wioland, F. 1983. La rythmique du français parlé, Strasbourg : Institut international d'études françaises.

WinPitch. 1995-2020. Speech analysis software, www.winpitch.com.

Practical challenges in polylexical metronome synchronisation

Connor McCabe
Phonetics and Speech Laboratory, University of Dublin, Ireland
https://doi.org/10.36505/ExLing-2020/11/0031/000446

Abstract

A metronome synchronisation paradigm was used to investigate possible interaction between P(erceptual)-centre location and initial mutation in Irish (Gaelic). A pilot study returned promising results in a subset of stimuli. Data later collected using only these stimuli differed from those of the pilot: participants exclusively synchronised metronome beats with the possessive particle preceding main lexical targets. In the pilot task, monosyllables as the minimal target, as opposed to exclusively disyllabic two-word possessive phrases in the modified task, may have focussed participants on these. This conditioning has been reinstated in a third version of the task by preceding target phrases with a dummy monosyllable.

Keywords: Methodology, P-centre, metronome synchronisation

Introduction and background

This paper discusses a problem arising with participant treatment of a metronome synchronisation task. Relevant context and background are presented briefly below. The problem itself is then detailed, and its speculated provenance is discussed. A solution to the issue for future work is sketched. This is followed by a summary and conclusion.

Background on P(erceptual)-centres and initial mutation in Irish

The present matter arose as part of doctoral research involving the relationship between prosody and morphophonology in Irish (Gaelic). Part of this work is concerned with *P-centres* (the perceptual 'downbeat' of a syllable; Morton *et al.* 1976) and one of two *initial mutations* which characterise Irish lexical morphology.

P-centres and metronome synchronisation

P-centres are the moment of perceptual occurrence of a syllable (Morton *et al.* 1976). There has been considerable investigation of acoustic correlates of the P-centre (Marcus 1981; Barbosa *et al.* 2005; Villing *et al.* 2011, *inter alia*) and environmental influences on its location (Marcus 1981; Fowler and Tassinary 1981; Cooper *et al.* 1986; Fox and Lehiste 1987; Šturm and Volín 2016). P-centre movement has been described in response to onset complexity (Fowler and Tassinary 1981), as well as number of syllables in a target (Lehiste 1987). There has also been limited investigation of morphophonological influences on

P-centre location. Franich (2018) reports a significant effect of item derivation status on P-centre location in Medʉmba, a Bantu language with morphological prenasalisation of initial consonants.

There are numerous methods for estimating P-centre location (for an overview of existing methods, see Villing *et al.* 2011). For investigation of higher-level influences on P-centre location *metronome synchronisation* is popular, as it is easy to implement and straightforward in analysis (Fowler and Tassinary 1981; Fox and Lehiste 1987; Šturm and Volín 2016; Franich 2018). In this method, participants are instructed to repeat target stimuli 'in time' with an audible metronome beat (MB); the P-centre is estimated as the point with which participants align MBs.

Initial mutation

Initial mutation refers to a family of phenomena which alter the segmental character of a lexical item's left edge for morphosyntactic purposes. Two such mutations exist in Irish: historically spirantising 'lenition' (or *séimhiú*) and historically nasalising 'eclipsis' (or *urú*). These are variously used to signal tense, phrasal relationships, and grammatical gender and number. Eclipsis was selected for study in the present work, as its character and implementation were more amenable to controls relevant to the phenomenon under examination.

The relationship between mutated lexical items and their 'underlying' forms is a matter of interest in the literature, both in terms of production (Falc'hun 1950; Scully 1973; Ball and Müller 1993; Welby *et al.* 2016) and perception (Usisshkin *et al.* 2017). This study investigated whether P-centre locations varied between mutated and unmutated members of homophone pairs (e.g. *a bpá* 'their pay' and *a bá* 'her sympathy', both /ə bˠɑː/).

Two metronome synchronisation tasks

A metronome synchronisation task was designed to investigate the relationship between eclipsis and P-centre location. In a pilot study, a large list of stimuli was prepared. Target monosyllables containing a long vowel were presented: in isolation (e.g. *pá* 'pay'), following the possessive particle *a* with or without eclipsis (e.g. *a* (*b*)*pá* 'her/their pay'), or in one of two three-word prepositional phrases (e.g. *ar a* (*b*)*pá* 'on their/her pay' or *ar an bpá* 'on the pay'). Additionally, target lexical items could have either a singleton or cluster onset. Each of these conditions had a particular reason for inclusion at the time of design, relating to previous work on P-centres.

Three native speakers of the Waterford sub-variety of Munster Irish were recruited for the pilot study. Stimuli were presented on a desktop monitor while participants heard a 60 beats-per-minute MB over headphones. When a new stimulus was presented, they were instructed to repeat the word/phrase 'in time' with the MB until it was changed. Two of three participants treated the

task in the intended way. Results for these two participants showed that, for simple possessive phrases with singleton-onset main lexical targets, estimated P-centre location within onsets corresponded strikingly between lexical matches (e.g. *a pá* 'her pay' – *a bpá* 'their pay'), rather than between homophones (e.g. *a bá* 'her sympathy' – *a bpá* 'their pay'). The third participant did not show this correspondence, but did robustly differentiate between mutated items and their unmutated homophones.

For main data collection, the metronome task itself was unchanged. Stimuli, however, were now restricted to simple possessive phrases with singleton onsets. These represent the overlap of the simplest environment for eclipsis (possession) and maximal segmental control (possessive *a* + bilabial consonant + /ɑ:/), allowing presence/absence of mutation to be the only substantial variable. 7 native speakers of Waterford Irish took part in the first round of data collection. Instructions were modified to prevent the task interpretation shown by the third pilot participant.

Results were extremely uniform, but completely unlike those of the pilot. Instead of showing any clustering of P-centre location by condition (whether by lexical {***p**á* – *b**p**á*, ***b**á* - *m**b**á*} or homophonic {***b**á* – ***bp**á*, ***mb**á* - ***m**eá*} pairing), MBs consistently fell roughly 150ms before target lexical item onset. This was not merely inconvenient, it was inconsistent with the psychoacoustic basis of the P-centre. P-centres are well-documented to move as a function of onset characteristics. Participants seem to have synchronised MBs not with the target lexical items, but with the initial possessive particle *a*. The latter's invariance across all conditions is consistent with the uniform placement of MBs. In the pilot study, the inclusion of target items in isolation, alongside phrasal contexts, may have unintentionally, but helpfully, guided participants' attention.

Going forward, all target phrases will be preceded by a dummy monosyllable *tá* to be repeated in alternation with the target itself (e.g *tá – a pá*). This will serve to reinforce participant focus on the lexical items of interest, rather than on whole phrases with initial *a* as synchronisation target.

Conclusion

In the absence of guidance, participants in a metronome synchronisation task may align MBs with the wrong part of a stimulus. This is a challenge for the collection of data on P-centre location in polylexical contexts, as in the study of Irish initial mutation described above, the nature of which required the use of phrasal stimuli with the items of interest in non-initial position.

References

Ball, M., Müller, N.. 1993. *Mutation in Welsh*. London & New York, Routledge.

Barbosa, P.A., Arantes, P., Meireles, A.R., Vieira, J.M. 2005. Abstractness in Speech-Metronome Synchronisation: P-Centres as Cyclic Attractors. Proc. Interspeech 2005, 1441-1444. Lisbon, Portugal.

Cooper, A.M., Whalen, D.H., Fowler, C.A. 1986. P-centers are unaffected by phonetic categorization. Perception & Psychophysics 39, 187-196.

Falc'hun, A.F. Le system consonantique du Breton avec une étude comparative de phonétique experimentale. Annales de Bretagne 57(1), 5-194.

Fowler, C.A., Tassinary, L. 1981. Natural measurement criteria for speech: The asynchrony illusion. In Long, J., Baddeley, A. (eds), Attention and performance (Vol. IX), 521-535. New Jersey, Erlbaum.

Franich, K. 2018. Tonal and morphophonological effects on the location of perceptual centres (p-centers): Evidence from a Bantu language. Journal of Phonetics 67, 21-33.

Fox, R.A., Lehiste, I. 1987. Effect of unstressed affixes on stress-beat location in speech production and perception. Perceptual and motor skills 65(1), 35-44.

Morton, J., Marcus, S., Frankish, C. 1976. Perceptual Centers (P-centers). Psychological Review 83(5), 405-408.

Scully, C. An experimental study of Welsh nasal mutations. Phonetics Department Report 4, 47-70. University of Leeds.

Šturm, P., Volín, J. 2016. P-centres in natural disyllabic Czech words in a large-scale speech-metronome synchronization experiment. Journal of Phonetics 55, 38-52.

Usisshkin, A., Warner, N., Clayton, I., Brenner, D., Carnie, A., Hammond, M., Fisher, M. 2017. Lexical representation and processing of word-initial morphological alternations: Scottish Gaelic mutation. Laboratory Phonology 8(1), 1-34.

Villing, R.C., Repp, B.H., Ward, T.E. 2011. Measuring perceptual centers using the phase correction response. Attention, Perception & Psychophysics 73, 1614-1629.

Welby, P., Ní Chiosáin, M., Ó Raghallaigh, B. 2011. A phonetic investigation of Irish eclipsis: Preliminary results and challenges. Proc. ICPhS XVII, 2122-2126. Hong Kong, China.

Welby, P., Ní Chiosáin, M., Ó Raghallaigh, B. 2016. A total eclipse of the heart? The production of eclipsis in two speaking styles of Irish. Journal of the IPA 47(2), 125-153.

The production of Italian dental affricates by Portuguese speakers

Chiara Meluzzi
Department of Linguistics, University of Pavia, Italy
https://doi.org/10.36505/ExLing-2020/11/0032/000447

Abstract

This work deals with the production of Italian dental affricates /ts dz/ by 2 female Portuguese speakers. Due to the lack of affricates in their L1, the aim was testing whether an affricate articulation and lengthening is preserved across phonological contexts. Through sentence-list reading, it will be shown how the affricate articulation is generally preserved, although /dz/ tends to reduce into a fricative /z/ more than the voiceless /ts/. An intermediate degree of sonority has also been detected together with a peculiar lengthening of the affricate involving only the occlusive part, whereas the fricative remains unvaried. These results may indicate a different phonological status of dental affricates for LS speakers.

Keywords: dental affricates, language acquisition, phonetics, Italian.

Italian dental affricates: L1 and LS

Dental affricates are considered typologically rare and marked sounds (Maddieson 1984). These sounds are difficult to be perfectly acquired by L2 or LS speakers (cf. Costamagna 2008, Sorianello 2019). In particular, the voiced /dz/ tends to be substituted with the fricative counterpart /z/ and/or to loss voicing, partially or completely. This also reflects the evolution of voiced affricates in the phonological repertoires of many languages (Żygis 2008: 23).

Italian has preserved a phonological opposition between voiceless and voiced dental affricates, but with a great variation within Italian dialects (cf. Meluzzi 2020: 43). Detailed phonetic analysis of Italian L1 speech has highlighted the emergence of an intermediate sonority degree, in which voicing is only present in the occlusive part. The emergence of this third degree of sonority has been identified in Italian native speakers in particular sociophonetic settings, such as in dialect-contact situations (cf. Meluzzi 2020). Conversely, Portuguese phonology does not present affricates, with the partial exception of post-alveolars in loanwords.

This work offers a preliminary analysis of the phonetic variation in terms of voicing and lengthening of dental affricates as produced by two Portuguese speakers. The hypothesis is that these sounds will be reduced into fricatives, and that they will differ from L1 in both voicing and lengthening.

Methodology

Two 35 y.o. female speakers have been recorded in a sound-proof room at University of Pavia in different moments. They came from the same town in Portugal, and were living in Northern Italy since 5 years at the moment of recordings. They did not receive a formal education in Italian, although they engage regularly with Italian friends.

Speakers were asked to read twice a list of 52 Italian realistic sentences, each one containing a real word with a dental affricate in different phonological constraints, and an equal number of distractors. The sentences were all declaratives with a balanced prosodic contour (e.g., C'è un pazzo sulla strada, "There's a mad man on the road"). Among the 52 stimuli, 18 tokens presented the dental affricate in intervocalic geminated contest, 8 in singleton, 10 as word-initial and 16 after a sonorant /l/-/n/-/r/. The distribution of tokens aimed at mirroring the phonological distribution of dental affricates in the Italian lexicon. The total 212 tokens were manually annotated in PRAAT following Meluzzi's (2016, 2020) protocol. Affricates were classified as either voiced, voiceless or intermediate, according to the presence of the voicing bar; the presence of a gap between the occlusive and fricative part was also annotated.

Analysis

The analysis was conducted from both a qualitative and a quantitative perspective. Qualitatively, different phenomena have been noted such as: (1) difficult in maintaining vocal folds vibration through the whole phone; (2) presence of a gap between the occlusive and the fricative portion; (3) different length of the occlusive and the fricative portion across phonological contexts.
Affricates tends to be reduced to fricatives in about the 30% of the cases (20 cases for /ts/>/s/, 42 for /dz/>/z/); this frication is not determined by the phonological context. As in native speech, the distribution of sonority of dental affricates is determined by the phonological context (cf. Fig. 1): intermediate realizations appear more in intervocalic contexts, both geminate and singleton. Finally, voiced affricates in post-sonorant and geminate contexts also present a partial limitation of vocal fold vibration.

A difference in duration of the dental affricates and of the occlusive portion has also been found as determined by the phonological context. Conversely, the fricative portion of the affricate stays unvaried across contexts. Intermediate affricates do not vary across phonological contexts.

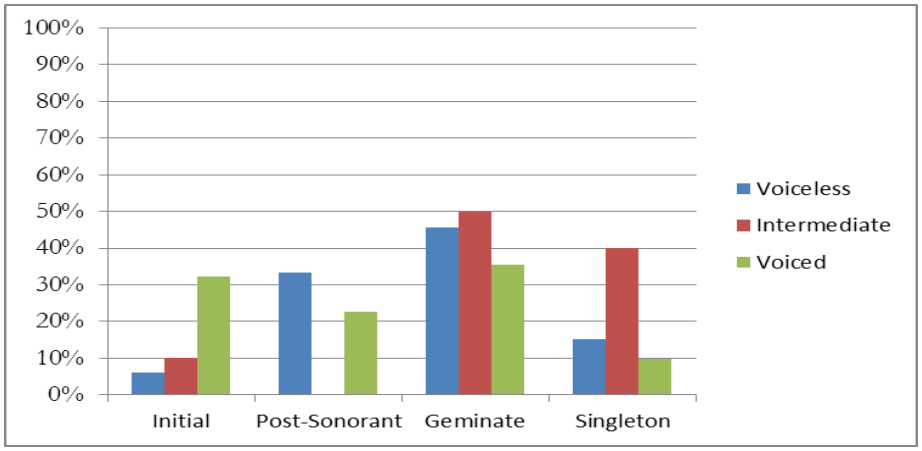


Figure 1. The realization of dental affricates as voiced, voiceless or intermediate according to the phonological contest ($\chi^2(6)=15{,}083$, p=,02, Cramer's V=,319)

Table 1. Duration of voiced and voiceless dental affricates across phonological contexts (ANOVA: /ts/ F(3,29)=7,612, p=,001; /dz/ F(3,29)=2,721, p=,006).

	Initial	Post-Sonorant	Geminate	Singleton
/ts/	185,75 ms. (st. dev. 10,8)	134,62 ms. (st. dev. 4,58)	212,51 ms. (st. dev. 4,37)	230,16 ms. (st. dev.27,77)
/dz/	136,86 ms. (st. dev. 5,06)	96,38 ms. (st. dev. 4,32)	166,03 ms. (st. dev. 5,86)	160,03 ms. (st. dev. 4,98)

Tab. 2. Duration of the closing portion of voiced and voiceless dental affricates across phonological contexts (ANOVA: /ts/ F(3,29)=6,497, p=,002; /dz/ F(3,29)=7,802, p=,001).

	Initial	Post-Sonorant	Geminate	Singleton
/ts/	72,47 ms. (st. dev. 2,8)	34,27 ms. (st. dev. 1,51)	74,49 ms. (st. dev. 2,92)	64,9 ms. (st. dev. 1,69)
/dz/	51,55 ms. (st. dev. 2,18)	21,14 ms. (st. dev. 7,47)	64,98 ms. (st. dev. 2,81)	46,11 ms. (st. dev. 1,05)

The data in Tab. 1 shows how voiced affricates are always shorter than voiceless ones. No great difference occurs between the two intervocalic contexts, whereas affricates are sensibly shorter in post-sonorant context, also confirmed by the post-hoc Tukey test. In voiced affricates, the occlusive portion is shorter in singleton and post-sonorant contexts as opposed to geminate one. For the voiceless affricates, however, the data confirms a similarity between the two intervocalic contexts vs. the post-sonorant one.

Discussion

Contrary to previous study, our data shows that non-native Italian speakers reduces affricates into fricatives less frequently than expected, although a major difficulty is attested with the voiced affricate /dz/. The emergence of an intermediate sonority degree as also been attested. The similarity in durations of singleton and geminate affricates mirror the productions of native speakers, in which a general opposition has been found between intervocalic vs. post-sonorant contexts. What appears to be typical of non-native speech is that the lengthening processes affect only the occlusive part of the affricate, leaving the fricative unvaried in duration across contexts.

These data could indicate that for non-native speakers the affricate articulation is differently interpreted at a phonological level: more than a proper phoneme, affricates could have been interpreted as a sequence of an occlusive and a fricative. This explanation could also account for the presence, albeit randomly attested in our data, of a gap between the closure and the frication.

Conclusions and further perspectives

This preliminary study on the production of Italian dental affricates by Portuguese speakers has confirmed how dental affricates are complex sounds, easily to be acquired by non-native speakers. The different lengthening patterns highlighted indicated that the phonological status of dental affricates may be different in L2/LS with respect to Italian L1. Due to the limits of our speech sample, it appears evident that further research on the phonetic characteristics of Italian dental affricates are needed in order to ascertain their phonological status in particular in L2 speech. It will also be interesting to correlate different phonetic features with the perception of a "foreign" accent.

References

Costamagna, L. 2007. The acquisition of Italian L2 Affricates: the Case of a Brazilian Learner. Proceedings of the congress 'New sounds', 26-29.

Maddieson, I. 1984. Patterns of Sounds. Cambridge: Cambridge University Press.

Meluzzi, C. 2016. A New Sonority Degree in the Realization of Dental Affricates /ts dz/ in Italian. In Ball, M.J. & Müller, N. (eds.), 2016, Challenging Sonority. Cross-linguistic Evidence, 252-275. Lodon: Equinox.

Meluzzi, C. 2020. Sociofonetica di una varietà di koinè. Le affricate dentali nell'italiano di Bolzano. Milano, Franco Angeli. Open Access: http://ojs.francoangeli.it/_omp/index.php/oa/catalog/book/459

Sorianello, P. 2019. Le consonanti affricate: strategie di acquisizione in italiano L2, Italiano LinguaDue 11.1, 1-21.

Żygis, M. 2008. On the Avoidance of Voiced Sibilant Affricates. ZAS Papers in Linguistics 49, 23-45.

Using uncertainty for multi-domain text classification

Kourosh Meshgi, Maryam Sadat Mirzaei
[1]RIKEN Center for Advanced Intelligence Project (AIP), Japan
https://doi.org/10.36505/ExLing-2020/11/0033/000448

Abstract

Multi-domain learning allows for joint feature detection to promote the performance on a learning task. The shared feature space, however, has limited capacity and should include only the most discriminative task-independent features that are useful for all the tasks. To this end, we proposed a global-local task uncertainty measure to monitor the usefulness of features for all tasks, increasing their effectiveness and generalizability while disentangling them from task-specific features that are not helpful for other tasks. Besides, this measure can utilize unlabeled domain data, tapping the vast reserves of unlabeled data to have even better features. An experiment on a multi-domain text classification shows that the proposed method consistently improves the baseline's performance and improves the knowledge transfer of learned features to unseen data.

Keywords: Multi-Domain Learning, Uncertainty, Feature Disentanglement

Introduction

There are numerous applications for text classification, ranging from document categorization to author identification and sentiment analysis. While different texts may include different vocabulary, grammar, writing styles, and collocations, there are some intrinsic commonalities and potential similarities among them that make multitask learning (MTL) a suitable tool to tackle this problem. MTL exploits such commonalities and tries to learn a shared feature space that improves the accuracy and generalization of all text domains simultaneously. To do this, MTL shares some learned features from one task to another that might need it but has no dedicated data to learn it. It also focuses on learning only relevant features that can be useful for all tasks. It uses other tasks as a regularizer for one, that empowers the tasks to tackle noisy or high dimensional data. It also shifts the features toward generalization since the emergent features should be useful for other tasks and domains (Ruder, 2017). However, if a classifier only employs generic features, its performance on the target task would be suboptimal (Liu, Qui, & Huang, 2017). On the other hand, fine-tuning each task diminishes others' performance or restricts the generalization of the shared features (Lee et al., 2017). To this end, the idea of shared-private MTL has been proposed in which, along with the shared feature space, each task has its own private feature space that can learn task-specific features and boost the task performance (Liu et al., 2017).

However, some task-specific features still emerge in the shared feature space that is not generalizable (thus, not useful for all tasks), wasting the capacity of the shared space to learn more desired features (Ganin & Lempitsky, 2015). It is potentially redundant to the features emerging in the private space of some tasks. To maximally improve the generalization of the shared feature space, the task-specific and task-independent features should be separated (i.e., disentangled), and feature redundancy should be omitted.

To this end, we propose the notion of global-local task uncertainty discrepancy, in which we monitor how private-only and shared-only features can classify a specific domain, measure the uncertainty of the labels given by each feature sets, compare them, and maximize the difference between the uncertainty of two classifiers (built by shared or private-only features) for each task. In another view, the MTL loss function is considering the label inaccuracy for each task and the amount of task uncertainty caused by shared features for each task (that can be explained with some other features emerging in private feature space). This will disentangle shared and private feature spaces, pull the useful features from private space to the shared, and expell the redundant ones.

Disentangling features using uncertainty

In a deep multitask network, several tasks $k=1,\ldots,K$ are jointly trained on their respective data D_k, to form a shared parameter space in which shared features $\boldsymbol{f}_S$ emerges. If all tasks only use $\boldsymbol{f}_S$ as the feature set, the setting is called fully-shared MTL. However, fully-shared settings are not capable of capturing the complicated relationship between tasks, and therefore, the emergent features have suboptimal performance. To this end, private-shared architecture allocates a private feature space $\boldsymbol{f}_P^{(k)}$ to each task k. In collaboration with shared features, private features capture more patterns emerging in each task, which promotes the performance of individual tasks. However, in this naïve form, there is no guarantee that the useful and generalizable features emerge in the shared space. Also, some features are better suited for one or a few tasks (task-specific features) that can emerge in shared feature space-wasting the capacity of this space, limiting the overall effectiveness of shared features and jeopardize the generalizability of these features to unseen tasks. Thus, the task-specific features should be forced to evolve only in the private space of tasks, while task-independent features should be encouraged to appear in shared space.

We propose a global-local uncertainty measure to see if a feature is useful for all tasks, or only a few. Instead of directly using training or development errors that are suboptimal indicators for learning progress (Kendall, Gal, and Cipolla 2018), we used task uncertainty to have a sense of how the MTL is affected by changing the shared feature space in a particular way. This change is typically a result of supervised training of task k, where the network parameters (including the corresponding private space and the shared space) are tuned to reduce the training loss of the task. Here, we calculate the uncertainty of each task using

only shared features (disabling private features) on its dev set, when a task changes the shared feature space. We accumulate these values to provide the global task uncertainty measure, L_{gu}. We also include the uncertainty of the current task (i.e., the task that is being trained) using both shared and private features as a local task uncertainty loss, L_{lu}. These uncertainty measure are added to the task loss, L_{task}, to provide the final loss for task k.

$$L^{(k)} = L_{task}^{(k)} + \lambda_1 L_{gu} + \lambda_2 L_{lu}^{(k)}$$

in which λ_1 and λ_2 are regularization coefficients, and L_{task} is the sum of loss of each task. Note that instead of dev set, L_{gu} can be calculated using unlabeled data as uncertainty calculation doesn't require label.

Experiments and results

To measure the effectiveness of the proposed uncertainty measure on feature disentanglement and on the performance of the classifier, we conduct two sets of experiments. In the first experiment (Table 1), we train our baseline text classifier on 16 different text domains and compare it with a single-task classifier and other MTL classifiers, and investigate the effect of proposed uncertainty (UPS-MTL) on the performance. We also used unlabeled data in UPS-MTL+ to demonstrate the power of including unlabeled data.

Table 1. The accuracy (%) of the model trained on 16 tasks, compared to its single task LSTM baseline and other MTL classifiers. (first, second, third)

DOMAIN	LSTM	MT-CNN	FS-MTL	SP-MTL	ASP-MTL	UPS-MTL	UPS-MTL+
BOOKS	79.5	84.5	82.5	81.2	84.0	86.3	86.8
ELECTRONICS	80.5	83.2	85.7	84.7	86.8	88.4	88.9
DVD	81.7	84.0	83.5	84.0	85.5	87.0	87.5
KITCHEN	78.0	83.2	86.0	85.2	86.2	89.3	89.6
APPAREL	83.2	83.7	84.5	86.5	87.0	87.8	88.1
CAMERA	85.2	86.0	86.5	88.0	89.2	89.4	89.9
HEALTH	84.5	87.2	88.0	87.2	88.2	90.1	90.5
MUSIC	76.7	83.7	81.2	83.0	82.5	84.4	84.9
TOYS	83.2	89.2	84.5	85.2	88.0	88.2	88.7
VIDEO	81.5	81.5	83.7	83.2	84.5	87.0	87.4
BABY	84.7	87.7	88.0	86.7	88.2	90.7	91.0
MAGAZINES	89.2	87.7	92.5	92.0	92.2	92.7	93.3
SOFTWARE	84.7	86.5	86.2	87.0	87.2	87.5	87.8
SPORTS	81.7	84.0	85.5	87.2	85.7	86.3	86.7
IMDB	81.7	86.2	82.5	84.7	85.5	85.7	86.3
MR	72.7	74.5	74.7	76.0	76.7	76.8	77.1

Table 2. The accuracy of the model trained on 15 domains and tested on the leave-one-out domain.

LOO DOMAIN	SP-MTL	ASP-MTL	UPS-MTL	LSTM-X
BOOKS	82.2	83.2	86.5	82.9
ELECTRONICS	84.7	82.2	86.3	83.3
DVD	85.2	85.5	86.4	84.0
KITCHEN	85.0	83.7	86.6	82.2
APPAREL	85.2	87.5	87.4	84.6
CAMERA	86.7	88.2	88.3	86.1
HEALTH	85.5	87.7	88.9	86.6
MUSIC	80.0	82.5	86.4	81.2
TOYS	86.2	87.0	87.2	84.3
VIDEO	85.7	85.2	86.7	83.5
BABY	83.5	86.5	86.6	85.4
MAGAZINES	89.5	91.2	91.1	89.8
SOFTWARE	87.0	85.5	87.3	85.6
SPORTS	83.7	86.7	86.9	83.7
IMDB	87.2	87.5	87.6	84.5
MR	74.0	75.2	75.4	74.0

In the second experiment (Table 2), we probe the effect of the proposed regularization on the generalization of the emergent shared features. In this experiment, we train the competing classifiers on 15 domains and test them on the remaining unseen domain. We also examine the viability of transfer learning by putting learned shared features of UPS-MTL in a single task LSTM-based

classifier (denoted by LSTM-X). The dataset includes 14 product review collections and two different movie review repositories for the task of binary classification similar to (Liu et al., 2017). Each category has a train/dev/test/unlabeled data size of approximately 1400/200/400/2000. We compared our method with a single task classifier (LSTM), PS-MTL, and Adversarial version of that APS-MTL proposed in (Liu et al., 2017), fully shared (FS) MTL, and also MT-CNN (Collobert & Weston 2008) with partially shared CNN for different tasks. The baseline classifier is an LSTM (128 units) operating on GloVe embedding of the input, followed by a dense and a Softmax layer. The regularization weight is tuned by cross-validation (λ_1=.01, λ_2=.025).

As Table 1 shows, the proposed regularization improves the performance of its baseline (PS-MTL) in almost all categories, and although not significant (p=*0.06*), it yields better performance compared to applying adversarial training on the same baseline. Adversarial training in APS-MTL aims to push out domain-specific features from shared space, which indirectly leads to better performance. However, in our proposed algorithm (UPS-MTL), the uncertainty regularization pushes out any features that increase global task uncertainty (punishes task-specific features) and, at the same time, reduces local task uncertainty (by encouraging each task to find better task-dependent features in their private space). The task-independent features in a shared layer of our proposed method are more generalizable compared to FS-, PS- and APS-MTL, as it is demonstrated by the second experiment (Table 2). Also, it is evident that the transferring emergent features in our model to a simple classifier without fine-tuning leads in an acceptable result (LSTM-X).

Conclusion

In this paper, we proposed a global/local task uncertainty regularization to disentangle task-independent and task-specific features in a private-shared MTL setting. Findings show that the features obtained by this measure improve the performance of MTL while providing a strong feature basis for transfer learning targeted at unseen domains.

References

Collobert, R.; and Weston, J. 2008. A unified architecture for natural language processing: Deep neural networks with multitask learning. Proc. ICML'08.

Ganin, Y., & Lempitsky, V. 2015. Unsupervised domain adaptation by backpropagation. Proc. ICML'15, 1180–1189, Lille, France.

Kendall, A.; Gal, Y.; and Cipolla, R. 2018. Multitask learning using uncertainty to weigh losses for scene geometry and semantics. Proc. CVPR'18, USA, 7482–7491.

Lee, S. W., Kim, J. H., Jun, J., Ha, J. W., & Zhang, B. T. 2017. Overcoming catastrophic forgetting by incremental moment matching. Proc. NIPS 2017, 4652-4662, USA.

Liu, P.; Qiu, X.; and Huang, X. 2017. Adversarial multitask learning for text classification. Proc. ACL'17, 1–10, Vancouver, Canada.

Ruder, S. 2017. An overview of multitask learning in deep neural networks. arXiv.

Perception of word-final inserted vowels and syllabicity in Italian

Veronica Miatto
Department of Linguistics, Stony Brook University, US
https://doi.org/10.36505/ExLing-2020/11/0034/000449

Abstract

Perceptual evidence is provided in favour of the non-syllabicity of word-final vocalic elements in consonant-final nonce words in Italian. These are optionally present after words ending in consonants, and their syllabicity status is debatable. In the experiment, speakers listened to stimuli of nonce words presenting variable duration of word-final schwas and judged whether the stimuli were monosyllabic or disyllabic. The results strongly suggest that speakers of Veneto Italian do not phonologically distinguish between nonce words that present a word-final schwa and those that do not. In fact, stimuli were generally judged as monosyllabic. The results of the study support previous research stating that word-final schwas in Italian are non-syllabic, phonetic vowels.

Keywords: epenthetic vowels, intrusive vowels, perception, Veneto Italian

Introduction

The present study addresses the ability of epenthetic vowels to be syllable nuclei. Specifically, perceptual-based evidence is used to determine whether Italian native speakers will perceive word-final inserted vowels as syllabic. Consonant-final words in Italian might be adapted with a word-final vocalic element, so that a word like 'jet' can be adapted as [dʒɛt] or as [dʒɛtːə], with the word-final inserted vowel preceded by lengthening of the consonant, but much variation is reported.

The phonological status of this inserted vowel is unclear. Some authors consider it a syllabic (canonically epenthetic) vowel that constitutes the nucleus of a new syllable and it is inserted to repair a violated structure (Bafile 2002 and later work, Passino 2008, Broniś 2016). On the other hand, other authors consider it a non-syllabic (intrusive) vowel, which does not form a new syllable and appears under prosodic and metrical pressures (Grice et al. 2015, Grice et al. 2018) or a vocalic element that is part of the release of the consonant (Repetti 2012, Miatto et al. 2019).

In particular, Miatto (2020) provides evidence from a production study that word-final inserted vowels in Italian are non-syllabic. While the duration measurements are somewhat ambiguous in their phonological underlying structure, repetition and place of articulation effects imply that the inserted

schwas are non-syllabic, and their presence is due to a need to aid the perceptibility of word-final consonants.

These results, compounded with the characteristics of the vocalic element itself, which satisfy Hall's (2006, 2011) diagnostics of intrusive vowels (vs. epenthetic vowels) in optionality, quality, and variability, strongly suggest that word-final schwas are non-syllabic. The present study provides additional, perceptual-based evidence that strengthens Miatto's (2020) conclusions.

Methodology

The study featured 13 native Italian speakers from Veneto (North-Eastern Italy), aged 19 to 36 years old.

In the experiment, they listened to 6 nonce words that followed the template C_1iC_2, such as /vik/. C_2 was a lengthened voiceless stop, followed by a word-final inserted schwa that would range from 0 to 100ms, in incremental steps of 25ms. No spelling or visual transcription was provided for the speakers, and the stimuli were produced by a native speaker of that variety of Italian. The stimuli were repeated in two blocks, in which the speakers had to judge whether the stimulus was constituted by one or two syllables.

Additionally, in the first block speakers were also asked to repeat aloud the stimuli they were hearing, to assess whether a stimulus presenting a word-final vocalic element would encourage the speakers to produce one too. The presence of a word-final schwa in the speakers' production was determined through spectrogram analyses.

Results

Results of 13 participants show that speakers overwhelmingly perceive words like [vik:] or [vik:ə] as one syllable. In fact, out of 773 tokens, 718 (93%) were perceived as one syllable, while 55 (7%) were perceived as two syllables.

Statistical analysis performed with a generalized linear mixed-effects model found that speakers are 3.15 times more likely to identify as disyllabic stimuli that end in coronals or labials than velars (95% confidence interval (CI) 1.04..9.53 times, $p = 0.0377$). In Miatto's study (2020), coronals and labials pattern together too, contrasting with velars. However, in that study, velars were less likely to be followed by a word-final schwa than labials or coronals, which would predict the opposite perceptual result to the one found in this study. While these opposing effects will have to be explained in future studies, I will note that both in the stimuli and in speakers' production, velars had abnormally long VOTs for Italian, reaching 70 ms, which might have influenced the perception of word-final schwas.

The duration of the word-final schwa has an almost significant effect. Speakers tend to be more likely to identify a stimulus as disyllabic as the duration of the schwa increases, especially between 50 and 100ms (CI 0.92..5.89 times, $p = 0.069$).

When asked to repeat the stimuli, speakers would sometimes insert word-final vocalic elements, even if the stimulus did not contain one itself and vice versa. They displayed much variation in their production of inserted schwas, presenting one between 0% and 97% of the tokens. Crucially, the correlation coefficient between their rates of vowel insertion and whether they were perceiving one or two syllables was very low, and non-significant (cor = 0.098, p = 0.06). This is depicted in Figure 1, with the blue line indicating percentages of disyllables identified (regardless of the presence or duration of the final schwa), and the green line indicating the percentages of vowel insertion produced.

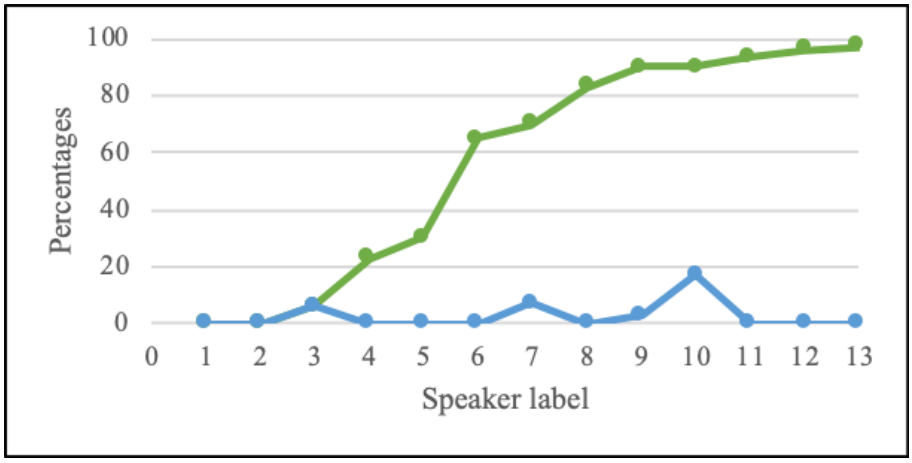


Figure 3. Correlation between percentages of disyllables perceived (blu) and word-final vocalic elements inserted (green) for each speaker

Some speakers would select some nonce words as disyllabic consistently, e.g. speaker S7 pressed the key for two syllables only for the nonce word [pit] (5/10 tokens), while speaker S3 perceived as disyllabic only the word [dit] and [kip] in some instances (3/10 and 2/10 respectively). This suggests that some speakers may have perceived some tokens as disyllabic because they had categorized that one particular lexeme as disyllabic, and not due to the presence of a word-final vocalic element.

Conclusions

The speakers' clear judgement of the stimuli as monosyllabic indicates that speakers cannot perceive the potentially syllabic difference between nonce words with an inserted vowel and nonce words that lack one. This conclusion is strengthened by the fact that the increasing duration of the word-final schwa was not a significant factor in the speakers' decisions, and the non-correlation

between the presence of the word-final vocalic element in the stimuli and whether the speakers would insert it in production.

In summary, the results of this study show that speakers are not able to discriminate between stimuli with a word-final schwa and stimuli that lack one. Both cases were mainly perceived as monosyllabic words, suggesting that the inserted vowel, when present, is a phonetic vowel and does not participate in syllabification processes, contrary to what was suggested in previous literature.

References

Bafile, L. 2002. Il trattamento delle consonanti finali nel Fiorentino: Aspetti fonetici. Atti delle XXIII Giornate G.F.S., Pisa, Italy.

Broniś, O. 2016. Italian vowel paragoge in loanword adaptation. Italian Journal of Linguistics 28, 25-68.

Grice, M., Savino, M., Caffò, A., Roettger, Tb. 2015. The tune drives the text - Schwa in consonant-final loan words in Italian. Proc. 18th International Congress of Phonetic Sciences (ICPhS 2015). Glasgow, UK.

Grice, M., Savino, M., Roettger, Tb. 2018. Word final schwa is driven by intonation-The case of Bari Italian. Journal of the Acoustics Society of America 143, 2474.

Hall, N. 2006. Cross-linguistic patterns of vowel intrusion. Phonology 23, 387-429.

Hall, N. 2011. Vowel epenthesis. In van Oostendorp, M., Ewen, Cj., Hume, E., Rice, K. (eds.). 2011. The Blackwell Companion to Phonology, 1576-1596. Oxford: Wiley-Blackwell.

Miatto, V., Hamann, S., Boersma, P. 2019. Self-reported L2 input predicts phonetic variation in the adaptation of English final consonants into Italian. Proceedings of Proc. 19th International Congress of Phonetic Sciences (ICPhS 2019). Melbourne, Australia.

Miatto, V. 2020. The syllabicity of paragogic vowels in Italian. (Submitted).

Passino, D. 2008. Aspects of consonantal lengthening in Italian. Padua: UniPress.

Repetti, L. 2012. Consonant-final loanwords and epenthetic vowels in Italian. Catalan Journal of Linguistics 11, 167-188.

A phonetic comparison of two Irish English varieties

Francesca Nicora[1], Sonia Cenceschi[2], Chiara Meluzzi[3]

[1]Department of Italian, National University of Ireland, Ireland

[2]Department of Innovative Technologies, University of Applied Sciences of Southern Switzerland

[3]Department of Linguistics, University of Pavia, Italy

https://doi.org/10.36505/ExLing-2020/11/0035/000450

Abstract

This research offers a preliminary survey on vowels and diphthong variation between two Irish English varieties: Galway (GW) and Letterkenny (LK). The results showed only a smaller difference between GW and LK with respect to the monophthongs, whereas a larger difference was found for the MOUTH diphthong. Despite the great amount of literature on English dialects, a phonetic investigation of these specific varieties is still lacking. This study may open the path to further investigations of sociophonetic values-and the stereotypes associated with different varieties, in particular those of the northern regions.

Keywords: Irish English, sociophonetics, MOUTH, vowels, phonetics.

Introduction

The past years have witnessed an increasing number of studies focused on the phonology of Irish English (IE), shedding light on new tendencies in IE spoken across the present-day Republic of Ireland. For example, Hickey (2005, 2007) has discussed the IE realisation of the MOUTH diphthong /au:/ in eastern dialects, highlighting that in Dublin the realisation begins from a front starting point of [æ], which contrasts with the traditional IE low starting point of [ɑʊ]. In present times, the rapid spread of new Dublin English across the country, especially among young females, is leading to a decrease of this feature that is traditionally associated with southern IE. On the other hand, previous research on the MOUTH diphthong showed that northern IE realisations differ from those in the south (Corrigan 2010). McCafferty (2007) specifically noted that no social or stylistic variation correlated to such variants were found while investigating the (London) Derry variety.

This research offers a preliminary survey on the cardinal vowels /a/ and /i/, and on the MOUTH diphthong variation across two IE varieties: Galway (GW) and Letterkenny (LK) that belong respectively to southern IE and northern IE. LK is located in a transition zone between the south and north, called County Donegal – a part of the Republic of Ireland where IE is of a broadly northern type. To the authors' knowledge, an investigation of the variety spoken in LK is

still lacking, and the number of studies for the Galwegian variety is also very limited (Peters 2016). A phonetic investigation and comparison between these two varieties is therefore considered desirable so that new data can be compared with previous literature. Our hypothesis is that a difference between the GW and LK spoken varieties is present, but that the GW variety was more prone to supra-regional southern Irish English. Furthermore, we supposed that the diphthong will show a greater variation between the two varieties, and that a stylistic variation may also be present.

Methodology

The subjects of this research are two 22-year-old adult females born and currently living in GW and in LK, respectively. Speech materials were drawn from the corpus of Irish English Speech (IES), which was designed to gather data on spoken speech of different Irish English varieties in the present-day Republic of Ireland (Nicora 2020b). Recording sessions recorded in a soundproof room followed the guidelines of the Interactive Atlas of Romance Intonation (IARI) project (Prieto et al. 2011-2014), and were based on the Discourse Completion Task (DCT; Blum-Kulka et al. 1989). The DCT was translated into English and adapted for Irish speakers as detailed by Nicora (2020a). Two target words, "mandarins" and "tangerines", and three words containing the MOUTH diphthong were selected for the analysis of the cardinal vowels /a/ and /i/ and the diphthong /ou/ respectively.

Analyses

142 tokens were manually annotated in PRAAT for the cardinal vowels /a/ and /i/ in the words "mandarins" and "tangerines" and for the MOUTH diphthong in the sentence "It's too loud, turn down the sound". A script was used to automatically extract the F0, F1, F2 and F3 values at 7 target points (dynamic analysis). A quantitative analysis was conducted only on monophthongs, as the number of stimuli containing the MOUTH diphthong did not allow for a proper quantitative analysis. However, due to the perceived variation in MOUTH between our two target varieties, we decided to first investigate these data qualitatively, before pursuing further analysis.

We performed a one-way ANOVA analysis on the cardinal vowels' formants. Data shown in Tab. 1 indicates that there is a significant difference only in F1 for both values given at the midpoint. Specifically, the LK F2 is significantly lower for /a/ than in GW, with a mean difference of approximately 163 Hz. Conversely, the first formant of /i/ is higher in LK than in GW. The ANOVAs show no significant results for /a/ F1, and for /i/ F2.

Table 1. Mean values of /a/ and /i/ vowels in GW and LK, with Anova results.

		Galway	Letterkenny	Anova	p value
/a/	F1	995.12 Hz (st. dev. 20.84)	957.515 Hz (st. dev. 52.29)	F(1. 51)=.652	=.423
	F2	1827.126 Hz (st. dev.) 17.31)	1663.85 Hz (st. dev. 23.52)	F(1.51)=28.535	<.0001***
/i/	F1	416.88 Hz (st. dev. 18.18)	482.8 Hz (st. dev. 19.89)	F(1.51)=5.039	=.030*
	F2	2493.83 Hz (st. dev. 15.09)	2295.38 Hz (st. dev. 10.61)	F(1.51)=12.02	=.102

No significant variation has been found between read and spontaneous speech. As for the MOUTH diphthong, the audio analysis has been supported by the qualitative investigation of a PRAAT spectrogram. In GW the diphthong is realised as [ɑʊ], while in LK it was realised as [øy], but with great within-speaker variation. In spontaneous speech, the diphthong in LK was reduced to a more centralized monophthong [ɵ] for about a quarter of the corpus. This is reflected in the variation of the F1 and F2 mean values, as calculated at the second and sixth time point, thus corresponding to the first and second element of the diphthong (Tab. 2). As is clearly seen from the values in Tab. 2, the MOUTH diphthong has been realised differently between the two varieties. Although auditorily perceived as [y], the second element of the diphthong also seems to differ between the two varieties, particularly concerning the F1, where we find lower values in LK than in GW.

Table 2. Mean values of the MOUTH diphthongs in GW and LK as measured at the second and sixth points in time.

	Galway [ɑʊ]		Letterkenny [øy]	
	2nd time point	6th time point	2nd time point	6th time point
F1	975.0 Hz (st. dev. 9.62)	962.25 Hz (st. dev. 7.84)	701.606 Hz (st. dev. 10.41)	513.918 Hz (st. dev. 11.03)
F2	1849.749 Hz (st. dev.) 13.79)	1665,048 Hz (st. dev. 14.59)	1454.978 Hz (st. dev. 9.81)	1771,613 Hz (st. dev. 10.47)

Discussion and conclusions

This preliminary investigation highlights the similarities and differences between the two Irish English varieties spoken in GW and LK. The data shows a tendency towards a centralization of the cardinal vowels /a/ and /i/, qualitatively confirmed for the MOUTH diphthong, which was also at times realised as a monophthong [ɵ] by the LK speaker. Both varieties maintain a low

starting point, which differs from the attested fronting in Dublin English (Hickey 2007). Furthermore, a lack of stylistic difference between read and spontaneous speech in the realisation of cardinal vowels has been noticed.

These preliminary results need to be confirmed by the collection of a wider sample and further analysis of stylistic variation, particularly of the northern varieties of Irish English. Further studies will be also necessary to test whether and how native speakers perceived this variation, especially in the MOUTH diphthong, and if there are any social values specifically associated with its reduction into a monophthong. The work confirms and updates what we knew about Irish English, and it opens the path to further investigation into the possible stereotypes associated with northern versus southern varieties.

Acknowledgements

The article has been written jointly by the authors. For the needs of the Italian Academy, F. Nicora wrote the introduction and methodology, and C. Meluzzi the analysis, discussion and conclusion. F. Nicora is responsible for the data collection, S. Cenceschi for the automatic extraction of values, and C. Meluzzi for the data annotation and statistical analysis.

References

Blum-Kulka, S., et al. 1989. Investigating cross-cultural pragmatics: An introductory overview. Cross-cultural pragmatics: Requests and apologies, 31, pp.1—34.

Hickey, R. 2005. Dublin English: evolution and change. John Benjamins Publishing.

Hickey, R. 2007. Irish English: History and present-day forms. Cambridge University Press.

McCafferty, K. 2007. Northern Irish English. In Language in the British Isles, D. Britain, (ed.), 122-134. Cambridge: CUP.

Nicora, F. 2020a, Prosodic training in foreign language: a study with Hiberno English learners of Italian, PhD thesis, unpublished.

Nicora F. 2020b. The corpus of Irish English Speech (IES), IRAAL conference, submitted.

Peters, A. 2016. Linguistic change in Galway City English. A variationist study of (th) and (dh) in urban western Irish English. Duisburg Papers on Research in Language and Culture, Vol.116. Frankfurt a.M.: Peter Lang.

Prieto, P., Borràs-Comes, J., Roseano, P. (Coords.) 2010-2014. Interactive Atlas of Romance Intonation. Web page: http://prosodia.upf.edu/iari/.

Production of Greek vowels by hearing-impaired children

Elina Nirgianaki[1, 2], Maria Bitzanaki[2]

[1]Department of Speech and Language Therapy, University of Peloponnese, Greece
[2]Department of Speech and Language Therapy, Metropolitan College, Greece

https://doi.org/10.36505/ExLing-2020/11/0036/000451

Abstract

The present study investigates the acoustic characteristics of Greek vowels produced by hearing-impaired children with profound prelingual hearing loss and cochlear implants. The results revealed a significant difference between vowels produced by hearing-impaired children and those produced by normal-hearing ones in terms of duration. Stressed vowels were significantly longer than non-stressed for both groups, while F0, F1 and F2 did not differ significantly between the two groups for any vowel, with the exception of /a/, which had significantly higher F1 when produced by hearing-impaired children. Acoustic vowel spaces were similar for the two groups but shifted towards higher frequencies in the low-high dimension and somehow reduced in the front-back dimension for the hearing-impaired group.

Keywords: hearing impairment, children, Greek vowels, duration, formants

Introduction

Hearing-impaired speech is characterized by deviations both in segmental and suprasegmental level that often affect intelligibility and communication. As far as the segmental level is concerned, several deviations have been observed not only in the production of consonants, but also of vowels, despite the fact that their production is considered much easier. Among others, vowel substitutions and neutralization have been frequently reported in the speech of hearing-impaired individuals due to limited auditory feedback (e.g. Ryalls et al. 2003).

Acoustically, vowels produced by hearing-impaired speakers have been revealed to cover more restricted formant space, with more overlaps than the ones produced by normal-hearing speakers (e.g. Ryalls et al. 2003). Moreover, excessive duration is among the suprasegmental errors that have been reported in several studies (e.g. Hasanvand, et al. 2017) for both adults and children with hearing impairment. However, the results regarding the duration in hearing-impaired children's speech remain contradictory, as there are also studies reporting no difference in comparison to normal-hearing children's speech (Ghayedlou et al. 2020).

With regards to the production of Greek vowels, Nicolaidis & Sfakianaki (2007, 2016) have found significant differences between vowels produced by normal-hearing and hearing-impaired adults, such as prolonged duration and reduced vowel space for the latter group. To our knowledge there are no

studies on vowel acoustic parameters of Greek hearing-impaired children's speech in the literature.

Along this line of research, the present study aims at investigating the acoustic characteristics of Greek vowels, and particularly, duration, F0, F1 and F2, produced by hearing-impaired children by comparing them with those produced by normal-hearing children.

Method

The speech material consisted of ten real CVCV words produced within the carrier phrase /'ipe __ 'pali/ ("He/She said __ again"). The first syllable was either stressed or non-stressed, containing the target vowel which varied among the five Greek vowels, while the preceding and following consonants were voiceless stops. Three hearing-impaired children aged 8-12 years old with profound prelingual hearing loss and cochlear implants and four normal-hearing peers produced the speech material three times at a normal tempo in a quiet room. A consent form was signed by the parents/carers. Duration, as well as F0, F1 and F2 (in the middle of each vowel) were measured in Praat software (Boersma & Weenink 2019). SPSS (v26) was used for data statistical analysis.

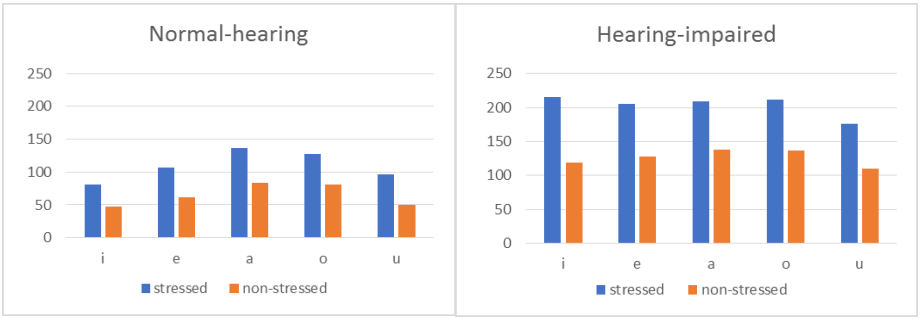
Results

Table 1 shows the mean values of duration, F0, F1 and F2 per vowel, for normal-hearing and hearing-impaired children, averaged across stressed and non-stressed vowels as well as across boys and girls.

Table 1. Mean duration (ms), F0, F1 and F2 (Hz) values per vowel for normal-hearing (NH) and hearing-impaired (HI) children.

Vowels	Duration		F0		F1		F2	
	NH	HI	NH	HI	NH	HI	NH	HI
i	64	167	257	262	401	408	2748	2618
e	84	166	243	247	667	615	2023	2035
a	110	173	255	244	832	931	1512	1551
o	104	174	241	260	621	591	1265	1400
u	73	143	235	255	438	393	930	951

A 4-way Anova (group x gender x stress x vowel) for each one of the dependent variables revealed a significant difference between vowels produced by hearing-impaired children and those produced by normal-hearing ones in terms of duration [$F(1,30)=44.207$, $p<0,0001$]. Subsequent t-tests per vowel revealed significant differences for all vowels (marginally for /a/). Stressed vowels were significantly longer than non-stressed for both normal-hearing [$t(38)=-5.061$, $p<0.0001$] and hearing-impaired children [$t(28)=-4.055$, $p<0.0001$] (see Figures 1, 2).



Figures 1, 2. The duration of Greek vowels for normal-hearing and hearing-impaired children in stressed and non-stressed condition, respectively.

F0, F1 and F2 did not differ significantly between the two groups for any vowel, except for /a/, which exhibited significantly higher F1 when produced by hearing-impaired children ($t(12)=-2.680$, $p<0.05$).

The main effect of vowel on F1 [$F(4,30)=41{,}737$, $p<0.0001$] and F2 [$F(4,30)=136{,}949$, $p<0.0001$], as well as subsequent post-hoc tests, indicated significantly distinguished vowel categories for the two groups in terms of both the low-high and the front-back dimension, respectively. In total, the acoustic space of hearing-impaired children is similar with the one of normal-hearing children but shifted towards higher frequencies in the low-high dimension and somehow reduced in the front-back dimension (Figure 3).

A main effect of gender was observed on F0 [$F(1,27)=8{,}808$, $p<0.05$], with boys having significantly higher mean F0 than girls (256 vs. 224 Hz).

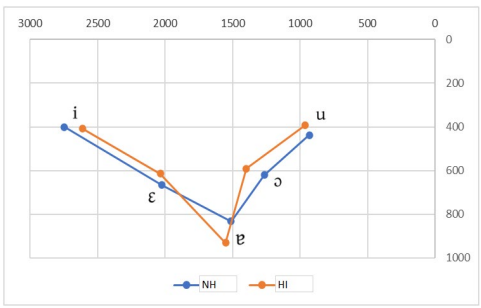


Figure 3. The acoustic spaces of Greek vowels as produced by normal-hearing (NH) and hearing-impaired (HI) children.

Discussion and conclusions

Overall, the results indicate significant temporal differences in vowel production between normal-hearing and hearing-impaired children and few differences in vowel acoustic spaces.

In terms of duration, vowels produced by hearing-impaired children were significantly prolonged, as previously reported for the English language too (Vandam et al. 2011). Vowels produced by normal-hearing children exhibited identical temporal hierarchical structure with those produced by both normal-hearing (Fourakis et al. 1999) and hearing-impaired adults (Nicolaidis & Sfakianaki, 2007), i.e. i<u<e<o<a; while, for hearing-impaired children the order was u<e,i< a,o, with all vowels having almost the same duration except for the shortest /u/. Stressed vowels were significantly longer than non-stressed for both groups, indicating that they both use the temporal parameters similarly to normal-hearing and hearing-impaired adults (Nicolaidis & Sfakianaki, 2007), in order to denote the effect of stress.

Regarding formant frequencies, vowel categories are clearly distinguished for all children. Compared to normal-hearing, hearing-impaired children show: (1) higher F1 for the low vowel, and (2) somehow reduced space in the front-back dimension; which may be related to the increased and reduced visibility of articulatory gestures forming the vowel constriction along the low-high and front-back axes of the vocal tract, respectively.

References

Boersma, P., Weenink, D. 2019. Praat: doing phonetics by computer [Computer program]. Version 6.0.50.

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

Ghayedlou, L., Ahmadi, A., Ghorbani, A., Torabinezhad, F., Keyani, M.R., Hashemi, S.B. 2020. Vowel duration measurement in shool-age children with cochlear implant. Int J of Pediatric Otorhinolaryngology, September, 136, pp.110-142.

Hasanvand, M., Torabinezhad, F., Abolghasemi, J., Eslami, M. 2018. A comparison of speech duration in cochlear implant and normal hearing elementary school students in the imitation and reading tasks. Func Disabil J, 1(1):21-30.

Nicolaidis, K., Sfakianaki, A. 2007. An acoustic analysis of vowels produced by Greek speakers with hearing impairment. ICPhS*,* August, vol. XVI, pp. 1969-1972.

Nicolaidis, K. & Sfakianaki, A. 2016. Acoustic characteristics of vowels produced by Greek intelligible speakers with profound hearing impairment II: The influence of stress and context. Int J Speech-Language Pathology, April, vol.18, no.4, pp. 1-14.

Ryalls, J., Larouche A., Giroux, F. 2003. Acoustic comparison of CV syllables in French–speaking children with normal hearing, moderate-to-severe and profound hearing impairment. Journal of Multilingual Communication Disorders 1, 99-114.

Vandam, M., Ide-Helvie, D., Moeller, M. P. 2011. Point vowel duration in children with hearing aids and cochlear implants at 4 and 5 years of age. Journal of Clinical Linguistics and Phonetics, August, vol.25, no.8, pp. 689-704.

Difficulties in adjacent vowel length of L1 Russian speakers in Czech

Zdena Palková, Tomáš Bořil, Jitka Veroňková
Institute of Phonetics, Faculty of Arts, Charles University, Czech Republic
https://doi.org/10.36505/ExLing-2020/11/0037/000452

Abstract

The sound category of the vowel quantity is applied in the structure of languages in different ways, and its adaptation from one system to another is difficult. The subject of the paper is the difficulties Russian speakers have in the production of Czech texts with more long vowels in a row, i.e., in a situation that does not exist in Russian. Sample of Czech created for the purpose of the experiment and recorded by Russian and Czech native speakers serve as the basis. The success in the realization of quantity in Russian speakers as assessed by Czech native listeners was monitored, and the duration values of short and long vowels and their ratio in the speech of Russian and Czech speakers were compared.

Keywords: Czech as L2, Russian as L1, vowel length, perception, word

Object of the research

The framework for comparing the language structures of individual languages is, at the phonological level, especially the relationship to meaning, and at the phonetic level, for example, the stability of individual characteristics in the competence of language users. It is useful to note the stability and changes in the treatment of a particular sound-relevant property in the course of transition from L1 to L2, and the way a native listener assesses the final realization in a foreigner's speech. (Major 2008, Colantoni et al 2015, 29–72)

In Czech, vowel quantity is a fundamental distinctive feature of the phonological system. The occurrence of long vowels is independent of word stress, the number of consecutive lengths and their word position. (Palková 1997) This structural independence causes difficulties for non-native speakers, including the advanced ones. In Russian, vowel length difference appears as a possible prosodic signal of the word stress. (Kasatkin 2006) In this paper, we examine how difficult it is for Russian speakers to observe the realization of the correct vowel quantity of two adjacent vowels, i.e., in a situation that does not exist in Russian, in a read Czech text.

Methodology

We recorded a text consisting of short sentences containing six-syllable sequences in the middle part. Within the sequences, the number of lengths (3 vs. 4 long vowels, 3L vs. 4L) and the word boundary (symmetrical 3syll|3syll

vs. asymmetrical 4syll|2syll variants) were controlled. Eight different structures were used, each represented by 3–4 carrier sentences: per speaker, there were 76 short (S) and 94 long (L) vowels, including 44 cases of adjacent long vowels (LL), available. The speaker group consists of 8 females with Russian as L1 (R) and 3 native Czech females (CZ) as a controlled group. Four phonetically educated native Czech listeners determined the acceptability of vowels in terms of their length. Using Praat software (Boersma, Weenink 2019), the acoustic analysis revealed the S and L's duration and their ratio.

Results

Results of the perception analysis

The success score of R in the realization of vowel quantity was 66%. The predominant error was shortening (92%; lengthening 8%). The target LL combination was successfully pronounced only in 33,6% of cases.

Variants 3L vs. 4L do not display any significant difference in the overall success rate (3L 34.1%, 4L 33.3%), nor does the difficulty of LL combination across the word boundary increase. The number of syllables in words seems to cause a slightly larger difference: the overall success rate of 3syll|3syll variants is 37.5% and the of 4syll|2syll variants is 29.2%.

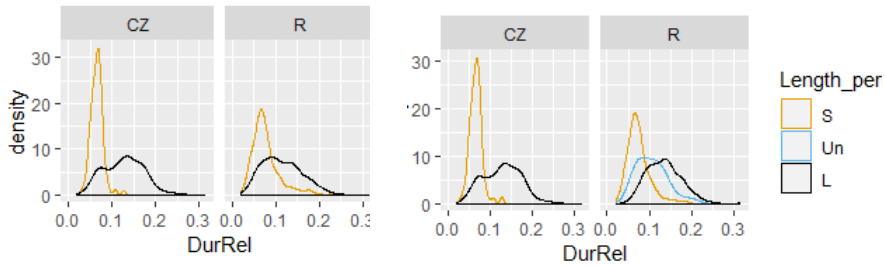
The differences may be clearly observed once the specific structures have been compared. Seemingly relevant is the S/L sequence in a word and the distinction in the symmetrical and asymmetrical distribution of syllables into two words. At the same time, both tendencies complement each other in our material. For example, the SLL sequence (success rate 40.1%) is preserved better than the SLLS sequence (14.3%), and the LL sequence in a two-syllable word is relatively successful (structure^{I} 44.4%, structure^{II} 37.5%). Comparing the structures with the same S/L sequences, a lowersuccess rate for the first word in the assymmetrical variant is always obtained (the most considerable difference in success rate is 6.2% (LLSS) compared to 38.1% (LLS) in the LLSSLL sequence).

Results of the acoustic analysis

In this section, the results of acoustic analysis of R in comparison to CZ are presented. Both the values of the distribution of normalized vowel durations and the values of duration ratio of long and short vowels were considered. The division of vowels into short and long was based on the original text, i.e., canonical form, and perceptual analysis, i.e., as perceived by native Czech listeners.

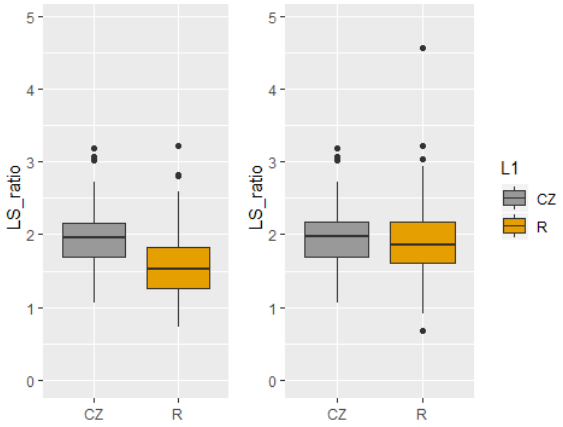
Figure 1 shows the distribution of vowel duration in L1 and L2 speakers. In CZ, S and L division is clearly perceived based on the original text, both in canonical length (a) and perception (c). In R, the duration distribution based on the original text is clearly differentiated only for S (b); L's duration with the highest frequency is objectively shorter than in CZ, comp.

(b) and (a). In the case of perception, a middle duration band is set aside, which is indifferent for the Czech listeners (d); S and L vowels have got visible peaks, even though the differences in the duration are lower than in CZ and quite variable, compare (d) and (a).



a b c d

Figure 1. Distribution of normalized vowel durations split according to the vowel length as written in the original text (a, b), and based on the perceptual analysis provided by Czech native listeners (c, d). S/L – short/long vowel, Un – unresolved vowel length; CZ – Czech speakers, R – Russian speakers.



a b

Figure 2. The ratio of long/short vowels as written in the original text (a) and based on a perceptual analysis (b). CZ – Czech speakers, R – Russian speakers.

Figure 2 shows the median of long and short vowel ratios. For text-based values (a), R's ratio is clearly lower than in CZ; the 50% interval of values overlaps only partially, and the median of R lies outside this interval. Duration ratio calculated according to native perception shows a more balanced result of both R and CZ groups; both groups approached the medians.

These results lead us to assume that native listeners project into the irregularities of R realizations a categorizing view, more phonological than phonetic.

Conclusion

a) To realize lengths in Czech, the determining unit for Russian speakers is the word and the word chain. The Russian speakers' intention seems to be to place and to realize long vowels on the right syllables. b) The Czech user probably relies more on the feeling of phonological opposition, i.e., contrast, than on the feeling of sufficiency in the duration of a particular sound. Creating such a relational basis in L2 production is difficult for the speaker, and the teaching requires a specially focused exercise. c) For the description of Czech, we obtained useful information about the acceptability of vowel duration in phonological length opposition.

Acknowledgements

This research was supported by the Czech Science Foundation project No. 18-18300S "Phonetic properties of Czech in non-native and native speakers' communication". We would like to thank Petra Poukarová and Martina Černá (Faculty of Arts, Charles University, Prague) to contribute to the listening analysis.

References

Boersma, P., Weenink, D. 2019. Praat: doing phonetics by computer [Computer program], version 6.0.25.

Colantoni, L., Steele J., Escudero, P. 2015. Second language speech. Theory and practice. Cambridge, Cambridge University Press.

Kasatkin, L. L. 2006. Sovremennyj russkyj jazyk. Fonetika. Moskva, Academa.

Major, Roy C. 2008. Transfer in second language phonology: A review. In Edwards, J. G. H., Zampini, M. L. (eds.), Phonology and Second Language Acquisition, 63–94. Amsterodam / Philadelphia, John Benjamins Publishing Company.

Palková, Z. 1997. Fonetika a fonologie češtiny. Praha, Karolinum.

Answering negative questions in Russian

Yulia Panchenko

Department of Theoretical and Applied Linguistics, Lomonosov Moscow State University, Russia

https://doi.org/10.36505/ExLing-2020/11/0038/000453

Abstract

In colloquial Russian answers "yes" and "no" to negative polar questions can either stand for confirmation or contradiction of the proposition of a question. This paper contains an experimental research of correlation between the choice of "yes" or "no" response particle and the way of expressing negation and the presence of *li* particle in a question and also of comprehension of the short "yes" answer. The results showed that there are several competing strategies for answering negative polar questions and acceptability judgments of possible answers in Russian, meanwhile the correlation with the presence of *li* particle in a question was approved.

Keywords: acceptability judgment task, negation, polar questions, yes/no answers, Russian

Introduction

Previous researches have shown that sentences which are both interrogative and negative are the most difficult to process, analyse and interpret (Savin, Perchonock 1965). Depending on the language negative polar questions (NPQ) may differ from positive polar questions (PPQ) by the response system (Evgrafova 1984), felicity conditions (Roelofsen et al. 2012) and ambiguity (Ladd 1981; Büring, Gungloson 2000; van Rooij, Šafářová 2003; Romero, Han 2004). Katz and Postal (1964) treat negative-question as a single feature, not a combination of the negative and the question.

In papers on negative polar questions in English, starting with (Ladd 1981) two types of NPQs are distinguished: with 'inside negation' reading or 'outside negation' reading. ON-NPQs have the same meaning as PPQs, the speaker wants a confirmation of positive proposition P (*You guys might be starving. You want to get something to eat? – Yeah, isn't there a vegetarian restaurant around here?*). These questions are relevant only when there is no evidence against P. IN-NPQs are used for confirmation of negative proposition -P and are relevant when there is a contextual evidence against P (*There's not really any place to go in Hyde Park – Oh, really, isn't there a vegetarian restaurant around here?*). It is possible to use the same terminology with Russian NPQs[1]. In this paper, only the ON-NPQs are considered.

NPQs also differ by the syntactic position of negation. Roelofsen et al. (2012) distinguishes low negation polar questions (*Did Lucy not go to Greece?*) and high negation polar questions with a negative clitic attached to the inverted

auxiliary (*Didn't Lucy go to Greece?*). In Russian, *ne*-questions (with the basic negative particle *ne* only) and *ne-li*-questions (with *ne* particle in combination with *li* particle and reversed word order) are quite similar to high and low NPQs. They slightly differ in thema-rhematic relations (Shatunovskii 2005) but often are interchangeable.

Negative polar questions might use a different response system from the positive ones. Evgrafova (1984) identifies two strategies for answering NPQs: (1) situational oriented – the *yes*-answer corresponds to a positive situation, the *no*-answer corresponds to a negative one; (2) communicative oriented – *yes*-answer corresponds to confirmation, *no*-answer corresponds to contradiction (some languages use special contradictional answers such as german *doch* or french *si*). The acceptance of short answers without the predicate repetition differs from one language to another. Languages can stick with one strategy or mix the both strategies in different ways:

3. English:
– Isn't it beautiful?
– No, it isn't. / Yes, it is.

4. Italian:
– *Non hai fame?* 'Aren't you hungry?'
– *No, non ho fame.* 'No, I'm not.' / *No, ho fame.* 'No, I am.'

5. Russian (colloquial):
– *Vy ne pomnite svoyu pervuyu knigu?* 'Don't you remember your first book?'
– *Net, ne pomnyu.* 'No, I don't remember.' / *Net, pomnyu.* 'No, I remember.' / *Da, ne pomnyu.* 'Yes, I don't remember.' / *Da, pomnyu.* 'Yes, I remember.'

As you can see, colloquial Russian allows to use any strategy for confirmation and negation (while the response system of standart language is closer to Italian).

Methodology

The aim of this paper is to analyze generation, processing and acceptability judgment of answers to negative polar questions in Russian depending on the way of expressing negation in a question (*ne-* and *ne-li* questions, negative predicatives and pronouns) with an acceptability judgment experiment.

Participants

The experiment was performed on 100 adult native Russian speakers (mean age 20.21, range 18–36).

Materials

Experimental materials included 20 Russian ON-NPQs, each one with 3 possible answers (response particle + predicate repetition): *yes*-agreement (negative confirmation: 'Yes, I don't remember), *yes*-denial and *no*-denial (positive contradiction: 'Yes, I do remember' or 'No, I do remember').

There were four types of questions: (1) NPQs with the basic negative particle, the *ne*-questions; (2) NPQs with negative predicatives (*net*, absence predicative; *nel'zya* 'impossible' etc.); (3) NPQs with negative pronouns or pronominal words; (4) NPQs containing the *li* particle, *ne-li*-questions.

Procedure

Participants were asked to evaluate each answer in terms of its grammatical acceptability using a Likert scale 1–5. Also, for each question they were asked to choose the most possible meaning of a short 'yes' answer (confirmation or contradiction).

Results

The results approved the hypothesis that *ne*- and *ne-li*-questions differ in their response systems. Meanwhile, the way of expressing negation in *ne*-questions doesn't matter. Acceptability judgments correspond with the short answer comprehension: in *ne*-questions the *yes*-agreement answer has a higher score on the acceptability and the preferred meaning is agreement, while in *ne-li*-questions the situation is reversed.

Table 1. Acceptability judgment task results.

	yes-agreement	*yes*-denial	*no*-denial	short *yes* meaning
ne-questions	3.84	3.03	3.64	agreement
negative predicatives	4.21	2.23	4.12	agreement
negative pronouns	3.85	2.49	3.4	agreement
ne-li-questions	2.54	4.18	2.69	denial

Discussion

According to the results, the response system for NPQs in colloquial Russian is freer than in the standard literary Russian language: situational and communicative oriented strategies are equally possible to use. It is interesting to compare response systems in colloquial and standart variants of other languages; for example, in English the situation seems very common, but it needs an experimental approval.

Notes

1. In Russian works these two types of NPQs have different names: questions with inherent or non-inherent thema (Baranov, Kobozeva 1983), interpretive or non-interpretive questions (Stepanova 1992, Dobrushina 2014), "negative" or "positive" negative polar questions (Shatunovskii 1980). These terms are equal to IN- and ON-NPQs.

References

Büring, D., Gungloson, C. 2000. Aren't positive and negative polar questions the same? California, UCSC/UCLA.

Dobrushina, E. R. 2014. What does the word *net* 'no' mean? In Dobrushina, E.R. 2014 Korpusnye issledovaniya po morfemnoi, grammaticheskoi i leksicheskoi semantike russkogo yazyka, 142-198. Moscow, PSTGU.

Evgrafova, S. M. On the issue of the affirmation and negation opposition. In Podlesskaya, V. I. (ed.), Lingvisticheskie issledovaniya: Tipologiya. Dialektologiya. Ehtimologiya. Komparativistika. Collected papers. Part 1, 97-107, Moscow, Nauka.

Katz, J. J., Postal, P. M. 1964. An integrated theory of linguistic descriptions. Cambridge, MIT Press.

Ladd, R. D. 1981. A first look at the Semantics and Pragmatics of Negative Questions and Tag Questions. Proceedings of Chicago Linguistic Society 17, 164-171.

Roelofsen, F., Venhuizen, N., Weidmann, S. 2012. Positive and negative polar questions in discourse. In Chemla E., Homer V., Winterstein G. (eds.) 2012 Proceedings of Sinn und Bedeutung 17, 455-472. Paris, ENS.

Romero, M., Han. C.-H. 2004. On negative yes/no questions. Linguistics and Philosophy 27(5), 609-658.

Savin, H. B., Perchonock, E. 1965 Grammatical structure and the immediate recall of English sentences. Journal of Verbal Learning and Verbal Behaviour 4(5), 348-353.

Shatunovskii, I. B. 2005. The main cognitive-communicative types of general questions in Russian. In Dialog: Komp'yuternaya lingvistika i intellektual'nye tekhnologii, 502-506, Moscow, Nauka.

Stepanova, E. B. 1992. *Ne*-particle in polar questions: meaning and scope. In Sistemnye semanticheskie svyazi yazykovykh edinits, 54-62, Moscow, MSU.

van Rooij, R., Šafářová, M. 2003. On polar questions. In Young R. B., Zhou Y. (eds.) 2003 Proceedings of Semantics and Linguistics Theory (SALT 13), 292-309. USA, CLC Publications.

Disfluency patterns in the language production system

Aurélie Pistono, Robert J. Hartsuiker
Department of experimental psychology, Ghent University, Belgium
https://doi.org/10.36505/ExLing-2020/11/0039/000454

Abstract

Within the language system, several of the language production levels may be involved in the production of disfluencies. Here, we conducted network task experiments to tackle disfluencies occurring during lexical selection, grammatical selection, and conceptual formulation. We showed that each difficulty induced a different pattern of disfluency. Additionally, multivariate pattern analyses demonstrated that difficulty is predictable from disfluency data patterns.

Keywords: language production, disfluency, multivariate pattern analyses

Introduction

Natural speech production is full of disfluencies. This term includes various phenomena such as filled or silent pauses, repeated words, and self-corrections. However, it is still not well understood why speakers are so often disfluent. To reveal the underlying cause of these phenomena, several authors attempted to relate the pattern of disfluencies to difficulties at specific levels of production, using network tasks (e.g., Hartsuiker and Notebaert, 2010; Figure 1). In this paradigm, participants describe a route through a network of pictures. This allows for the manipulation of the items to create difficulties at specific stages (e.g. conceptual generation) while holding others constant (e.g. lexical selection). Here, we conducted two experiments to examine the pattern of disfluency related to lexical selection difficulty, grammatical selection difficulty, and conceptual difficulty. Additionally, we aimed at examining whether, by contrast, the manipulated difficulty could be predicted based on the pattern of disfluency associated with it, using multivariate pattern analyses (MVPA, Haynes & Rees, 2006). Instead of analysing each dependent variable individually, MVPA extracts the information contained in the pattern of information available, to test whether experimental conditions can be distinguished from one another on the basis of the patterns observed.

Material and methods[1]

In each experiment, 20 bachelor students, all native Dutch speakers performed 20 network tasks. Pictures were either connected by one, two, or three straight lines or curves, that had a fixed length. The route through the network was indicated by a moving red dot that traversed the network in 42 seconds.

Instructions were given to provide an accurate description of the network using complete sentences and to synchronize the description with the dot that moved through the network.

In *Experiment 1*, we manipulated pictures' name agreement to examine the initial stage of lexical access. We also examined grammatical selection through grammatical gender. 160 pictures were selected from Severens et al., 2005 (i.e., eighty pictures had high name agreement and eighty had low name agreement; eighty pictures had a common gender name and eighty had a neuter gender name). Pictures were matched for their log frequency and age of acquisition. In *Experiment 2*, we examined the conceptual generation of the message, by manipulating the visual identification of some items (i.e., by blurring). We selected 160 pictures from the Multipic database (Duñabeitia et al., 2018), that we controlled for name agreement, age of acquisition, and visual complexity. For each participant, eighty pictures had a four pixel radial blur and eighty were control pictures.

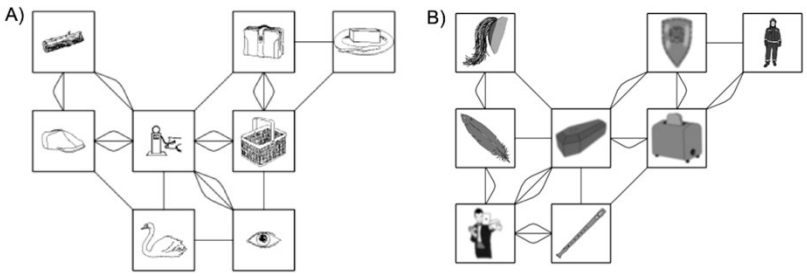


Figure 1. Example of a network for A) Experiment 1 and B) Experiment 2

All productions were transcribed and scored by a native Dutch speaker and checked by another native speaker. We analysed self-corrections, silent pauses, filled pauses, and prolongations related to each picture. In each experiment, we ran linear-mixed effects models with name agreement (low/high), grammatical gender (neuter/common), or conceptual difficulty (blurred/non-blurred items) as a fixed effect. Additionally, we used MVPA to predict whether each participant was about to mention a low or high name agreement item, a common or neuter gender item, or a blurred or non-blurred item. We trained a linear discriminant analysis classifier on the four disfluency features, using the Scikit-learn toolbox (Pedregosa et al., 2011).

Results

In Experiment 1, 16.3% of pictures induced at least one disfluency. Low name agreement items induced more self-corrections ($\chi^2(1)=10.19$, $p<.01$), silent pauses ($\chi^2(1)=17.59$, $p<.0001$), and filled pauses ($\chi^2(1)=11.32$, $p<.001$) than high name agreement. Common gender items elicited more prolongations

($\chi^2(1)=16.21$, $p<.0001$) than neuter gender items. The classifier could predict from the pattern of disfluency whether a participant was about to name low or high name agreement items (59.57% correct on average; $t(19)=7.04$, $p<.001$). The contribution of each feature was consistent across participants for this classification (self-corrections: ($t(19)=3.6$, $p<.01$); silent pauses: ($t(19)=6.5$, $p<.0001$); prolongations: ($t(19)=3.2$, $p<.01$); filled pauses: $t(19)=3.5$, $p<.01$). The classifier could also predict items' gender (53.72% correct on average; $t(19)=3.83$, $p=.001$). Only prolongations' contribution was consistent ($t(19)=-2.5$, $p<.05$).

Experiment 2 elicited more disfluencies than Experiment 1 (26% of pictures). Surprisingly, blurry pictures did not elicit more disfluency than control pictures. Classification accuracies were above chance level (53.24%; $t(18)=2.77$, $p<.01$) but the contribution of disfluencies was not consistent, suggesting inter-individual variability in disfluency patterns (Figure 2).

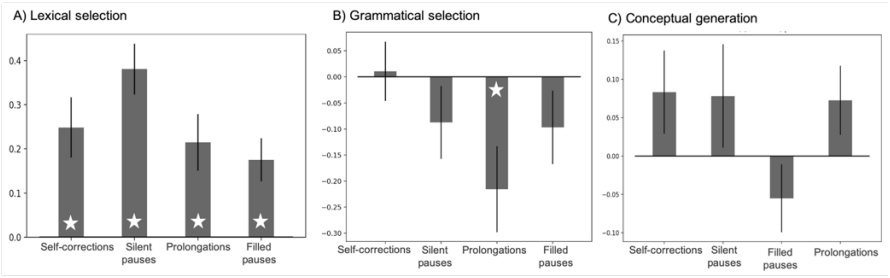


Figure 2. Contribution of each feature when classifying the pattern of disfluency related to each manipulation. White stars indicate significance.

Discussion

Difficulties in determiner selection led to a different pattern of disfluencies than difficulties in content word selection. This suggests that difficulties at distinct stages of production elicit different disfluencies. The finding that low name agreement induces pauses may indicate that they reflect an 'act of choice' between lexical items with similar semantic features (Beattie & Butterworth, 1979). Because of this 'act of choice', speakers are also more error-prone, leading to self-corrections. Surprisingly, common gender induced more disfluency than neuter gender. Possibly, this effect is related to the phonological form of this determiner ('de' in opposition to the neuter gender determiner 'het'), which is more likely to encourage prolongations.

In Experiment 2, impeding the visual identification of some items influenced the whole task, by increasing disfluency overall. It is possible that making the identification of some items more difficult might have hampered the monitoring system (i.e., poor error detection and correction while staying synchronized with the pace of the dot) throughout the whole task, similarly to time pressure (Oomen & Postma, 2001). MVPA provided complementary

findings. They revealed that the pattern of disfluencies is informative about the linguistic difficulty of an item, i.e. a classifier was able to learn and predict the type of item a speaker is about to name. For the conceptual manipulation in particular, it clarifies why linear mixed models were not able to reveal significant differences: conceptual difficuly manifests itself differently from one participant to another. These analyses represent a further step for current models of language production, to capture inter-individual variability.

Notes

1. Details about the project and methods are available here: https://osf.io/9yhcb/

Acknowledgements

This project has received funding from the European Union's Horizon 2020 programme under the Marie Sklodowska-Curie grant agreement No 832298.

References

Beattie, G.W., Butterworth, B.L. 1979. Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. Language and Speech, 22(39), 201–211.

Duñabeitia, J. et al. 2018. MultiPic: A standardized set of 750 drawings with norms for six European languages. Quarterly Journal of Experimental Psychology, 71(4), 808–816.

Hartsuiker, R.J., Notebaert, L. 2010. Lexical access problems lead to disfluencies in speech. In Experimental Psychology. 57(3), 169–177.

Haynes, J.-D., Rees, G. 2006. Decoding mental states from brain activity in humans. Nature Reviews Neuroscience, 7(7), 523–534.

Oomen, C.C., Postma, A. 2001. Effects of time pressure on mechanisms of speech production and self- monitoring. J Psycholinguist Res, 30(2), 163–184.

Pedregosa, F., et al. 2011. Scikit-Learn: Machine Learning in Python. J. Mach. Learn. Res., 12, 2825–2830.

Severens, E., et al. 2005. Timed picture naming norms for 590 pictures in Dutch. Acta Psychologica, 119(2), 159–187.

Degree of Czech R/L acquisition in L1 Chinese speakers

Petra Poukarová[1], Jitka Veroňková[2]
[1]Institute of the Czech National Corpus, Charles University, Czech Republic
[2]Institute of Phonetics, Charles University, Czech Republic
https://doi.org/10.36505/ExLing-2020/11/0040/000455

Abstract

This paper is based on the assumption that a successful acquisition of L2 is facilitated when the individual needs of particular students are taken into consideration. Our goal is to screen the perception and production of *r*- and *l*-sounds in Chinese students of Czech as L2. The experiment, in which 8 subjects participated, is based on minimal pairs and consists of a listening test (72 items) and recording of read speech (single words and sentences containing target words, 64 items per speaker). Within the set of minimal pairs, the position and the combination of the target sounds are controlled. The procedure was also tested regarding its suitability for the systematic screening of students' skills.

Keywords: Czech as L2, Chinese as L1, l-sound, r-sound, perception, production.

Introduction

It is well known fact that the Chinese have a problem distinguishing between r- and l-sounds when learning L2, including Czech, whose phonological system contains the alveolar trill /r/ and the alveolar lateral approximant /l/. These segments formminimal pairs and their incorrect pronunciation can cause problems in the intelligibility of spoken communication. We work with the assumption that mastering the production of sound phenomena should be supported by practicing perception (Strange & Shafer 2008), and that effective pronunciation training is based on the knowledge of the particular students' needs (Derwing and Munro 2015).

This paper describes the procedure used to screen the perception and production of Czech r- and l-sounds in Chinese students at the beginning of the phonetics course. The main goal was to obtain information about individual students and test whether the procedure would be suitable for a systematic use.

Methodology

8 speakers (7 females, 1 male) with Chinese as L1 participated in the experiments. They were students of a bachelor's degree of Czech Studies for Foreigners (Charles University, Prague) who had stayed in the Czech Republic for 3–4 months. Their level of Czech acquisition corresponded to B1–B2 of CEFR.

Experiment 1 – perception test

We formed a set of minimal pairs, which differed only in the target consonant *r* (R) and *l* (L). The following four positions were distinguished: (1) initial before vowel CV-, (2) intervocalic VCV, (3) final -VC and (4) two-consonantal cluster CC. Each position was represented by 3 pairs, i.e., 12 minimal pairs were at our disposal. The complete set was recorded by 3 native Czech females with precise pronunciation. The recordings served as the basis for a perception test designed using Praat multiple forced-choice environment (Boersma & Weenink 2019). The test contained 72 items altogether (i.e., 12 pairs x 3 speakers) randomly ordered; 5 more items were used in the training part. The subjects' task was to indicate which sound (R or L) they hear. The test was conducted individually or in a small group, the listeners used headphones.

Experiment 2 – production

The same set as in the Experiment 1 extended by 4 more CC pairs to make a record was used. First, subjects read the sentences, each of which contained one target word in the middle (32 sentences, 1 page). They were provided a little time to read the sentences in advance. Afterwards they read the whole set again as individual words (randomly ordered), without preparation. Recordings were taken individually in a sound-treated room (AKG C 4500 B-BC microphone, sample rate 48 kHz, 16-bit depth).

We calculated the number of correct answers (Experiment 1) and performed a perceptual analysis of the pronunciation of the target sounds R/L (Experiment 2). For perceptual analysis, we applied both a phonological view and monitored the phonetic quality.

Results

Experiment 1 – perception test

The ratio of incorrect items was 12% in both R and L, however, perception of each phoneme proved to cause difficulties in different patterns.

In the case of R, errors occurred predominantly in VCV. For example, the error rate for the lexeme *víry* (Eng. viruses), tested in 3 items, is 9 of 24 judgements, and for the lexeme *póry* (Eng. pores) 6/24. Compared to that, in the case of L, CC and CV- covered most of the errors. For example, the lexeme *blány* (Eng. membranes) showed an error rate 6/24, and the lexeme *luka* (Eng. meadows) 7/24.

The proportion of errors varied considerably between subjects: for 5 subjects, perception was easy, they had only up to 5% errors; 1 subject showed error rate of 12.5%; for 2 subjects, the task to discriminate *r*- and *l*-sound was difficult (error rate 30.6% and 38.9%).

Experiment 2 – production

511 of target items were analysed (255 within sentences – with 1 sample excluded because of a technical error, 256 single words).

71.2% of target items were pronounced with correct phonemes. Speakers were more successful in pronouncing the consonant L (83.9%) than the consonant R (58.6%). In this respect, the -VC position is specific: the number of correct Rs is higher than the number of correct Ls; the L items are often pronounced as labial approximants in this position.

In 16.8% of items, the original sound was substituted by the R/L counterpart. Substitution of R by L occurred with considerably higher frequency than, conversely, L by R (87.2% vs. 12.8%), for example *valu* (Eng. earthwork, Gen. sg) instead of *varu* (Eng. boil, Gen. sg) or *dluhy* (Eng. debts) instead of *druhy* (Eng. kinds). In one case, the speaker pronounced trill fricative /r̝/ instead of /r/.

Relatively often, in 11.7% of cases, realization was not intelligible enough to classify the sound as R or L unmistakably, mainly due to glottalization at the end of the word.

In general, success in production is similar regardless of whether the target items are tested as individual words or in the carrier sentence (see Table 1). The differences may be clearly observed once the specific positions have been compared. For example, the increase of the L to R substitutions in the CV-position in the sentences compared to the single words is apparent. The proportion of correctly pronounced Rs rises in VCV in the case of single words.

Table 1. Success rate of pronunciation r- and l-sounds (R, L) in single words and in the middle of a sentence. The table shows the rate score of phonologically correct pronunciation, the number of realizations the sound was substituted by the R/L counterpart, or could not be unambiguously classified.

	Words (N=256)			Sentences (N=255)		
%	L	R	Sum	L	R	Sum
Correct pronun.	41.0	30.5	71.5	42.7	28.2	71.0
Substitution R/L	13.7	1.6	15.2	15.7	2.7	18.4
Not clear			13.3			10.2
Other						0.4
			100.0			100.0

Regarding the subjects, there are distinct differences in the success rate (from 38% to 95% of correctly pronounced items). The three subjects with the lowest success in perception are among the four speakers with the lowest scores in production (the success rate up to 50.0%). On the other hand, the pronunciation of the three subjects was more than 90% successful. It is

possible to trace individual tendencies, concerning, for example, phonetic realization and positions.

Conclusion

The procedure proved to be feasible as it provides data on the perception and production of individual students. For both kinds of skill, it is useful to examine different positions of the target segments. The production of words as both single items and constituents of sentences can also display differences, so it is desirable to keep both types of text. The design of the listening test has also proved successful and its adaptation can be used to practice perception. The information obtained can contribute not only to the subsequent creation or selection of appropriate exercises, but also to a better understanding of the process of acquisition of the Czech sound patterns in Chinese speakers.

Acknowledgments

This research was supported by the Czech Science Foundation project No. 18-18300S "Phonetic properties of Czech in non-native and native speakers' communication". We would like to thank Tomáš Bořil (Institute of Phonetics, Charles University, Prague) for designing the perception test.

References

Boersma, P., Weenink, D. 2019. Praat: doing phonetics by computer [Computer program], version 6.0.25.

Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR), https://www.coe.int/en/web/common-european-framework-reference-languages.

Derwing, T. M., Munro, M. J. 2015. Pronunciation Fundamentals. Evidence-based Perspectives for L2 Teaching and Research. Amsterodam / Philadelphia, John Benjamins Publishing Company.

Strange, W., Shafer, V. L. 2008. Speech perception in second language learners: The re-education of selective perception. In Edwards, J. G. H., Zampini, M. L. (eds.), Phonology and Second Language Acquisition, 153–191. Amsterodam / Philadelphia, John Benjamins Publishing Company.

Preliminaries to the Tuvan interrogative intonation

Tatiana Ryzhikova, Albina Dobrinina, Ilia Plotnikov
Institute of Philology, Siberian Branch of RAS, Russia
https://doi.org/10.36505/ExLing-2020/11/0041/000456

Abstract

Tuvan is one of the minority Turkic languages of Siberia (RF). Its segmental structure is described quite well, whereas its suprasegmental level has not been studied yet. The purpose of this paper is to give a preliminary description of the Tuvan interrogative intonation. Several topical dialogues were recorded by three Tuvan women, yes/no questions were cut out and their structure and F0 changes were analyzed. The preliminary results show that Tuvan intonation seems to closely correlate with the information structure. The interrogative particle **be** appears to be pronounced with different F0 movement in accordance with the utterance focus. Generally Tuvan yes/no questions prove to be characterized by inclination of the tone, though the patterns are not as strict as in Altay, for example.

Keywords: intonation, yes/no questions, Tuvan, PRAAT, endangered languages

Introduction

The Tuvan language (formerly known as Uriankhay, Soyot, Tanu-Tuvan) is the language of the indigenous population of the Tyva Republic (RF). In Russia, it is also spoken in the South of the Krasnoyarsk region. The Tuvan ethnic group is settled in Mongolia and China – in the Xinjiang Uygur Autonomous region of China (Bavuu-Syuryun 2010: 58). The total number of speakers is about 270 thousand people (estimate of 2015), including about 253.7 thousand people (2010 census) in Russia, about 10 thousand – in Mongolia, and approx. 5 thousand – in China (Simchit 2017). 98% of Tuvinians living on the territory of Tuva consider Tuvan as a mother tongue (Bavuu-Syuryun 2010: 58).

The Tuvan language is of the agglutinative type and in accordance with the genealogical classification, it belongs to the Siberian group of the Turkic family of the Altai macrofamily of languages (Scherbak 1994: 39). In line with L. Johansons's classification it refers to the Northeastern (or Siberian), a southern, South Siberian, subbranch comprising Sayan Turkic: i.e. Tuvan, Tofan, Soyot, Dukhan, Tuhan, etc., and non-Sayan Turkic, i.e. Khakas, Shor, North Altay, Chulym, etc. (Johanson 1998). Tuvan has four dialects: Central, Western, North-Eastern (Todzhinsky) and South-Eastern. The central dialect is under consideration in this article.

If segmental structure of Tuvan is described quite well, its suprasegmental level has not been studied yet. Thus the purpose of this paper is to give a

preliminary description of the Tuvan intonation. The object of the study is the modal (yes/no) interrogative utterances of the native speakers of Tuvan.

Methodology

A topical dialogue was recorded by three Tuvan women. Then the interrogative utterances were cut out of the dialogue with the help of the Audacity program and their segmentation was done in PRAAT. The analysis was carried out at the syllabic level, because a syllable is the smallest unit of intonation.

The preliminary analysis of the material showed that additional micro-dialogues were necessary, thus the procedure was repeated with other sentences. Total amount of utterances under investigation is 74.

Results and discussion

Intonation plays one of the leading roles in establishing information structure. It seems to be especially true for Tuvan. Sh. Sat points out that there are four ways of making questions in this language: 1) using special adverbs (*where*? *from where*? etc.); 2) using special pronouns (*which*? *what*? *who*?); 3) using special interrogative particles (*be, ale (ele), aa, irge be, che, dee, ijik be*, etc.); 4) using special interrogative affix *-yl*. But the author also emphasizes that question can also be made just by using *special interrogative intonation* (Sat 1983: 33). Thus defining topic and focus in the utterance becomes of crucial importance for establishing Tuvinian intonation patters.

Discussed below are a few examples from Tuvan:

Figure 1 illustrates the utterances Hünüvüs dunup kel'di be? – Hünüvüs deep kel'di.

The question Hünüvüs dunup kel'di be? 'Has our sun risen?' starts from 14.4 st. increasing up to 19.6 st. on the last syllable of hünüvüs, then an even tone of about 18 st. is preserved on dunup followed by a slight decrease to 16 st. on the first syllable of kel'- and then again a sharp increase on -di up to 20 st. that is kept on 19, 5 st. on the particle be.

The response Hünüvüs deep kel'di. 'Our sun has risen' is pronounced as follows: it starts at 15 st. followed by a drop to 12 st. and a last sharp increase on the last syllable of hünüvüs up to 15 st., then again a drop to 11.5 st. on deep and a slight increase to 12 st. on kel'di.

It seems possible to conclude that the question starts either at the same level as a statement or even higher, but then the intonation deviates: the question has rising movement of F0, whereas a statement is characterized by declination. Other examples under consideration prove this conclusion to be true.

The utterance *Hüresh bolur be*? 'Will there be a traditional fight?' represents a canonical yes/no question with the interrogative particle **be** at the end and focus on the predicate *bolur* (Figure 2).

Generally, intonation is rising starting from the first word *hüresh* up to the last syllable of *bolur,* and *be* can be pronounced either with further inclination or with declination of F0.

The utterance *Üsh eves be*? 'Not three times?' (Figure 3) proves the idea of Sh. Sat that the utterance can be made interrogative just by the special intonation.

Here all the sentence is a focus and inclination starts at *üsh* and proceeds up to *be*, which can be pronounced either with a high tone or with a slight declination.

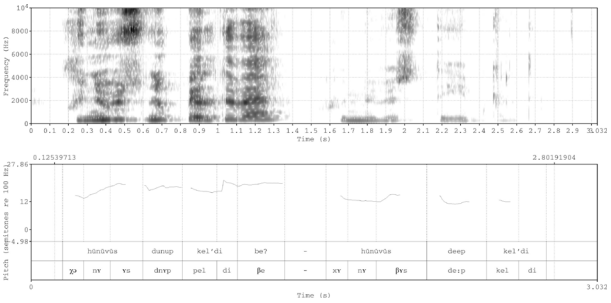


Figure 1. Intonogram of the utterance *Hünüvüs dunup kel'di be*? – *Hünüvüs deep kel'di.* 'Has our sun risen? – Our sun has risen.'

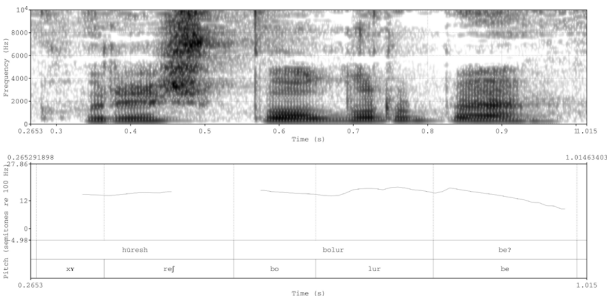


Figure 2. Intonogram of the utterance *Hüresh bolur be*? 'Will there be a traditional fight?'

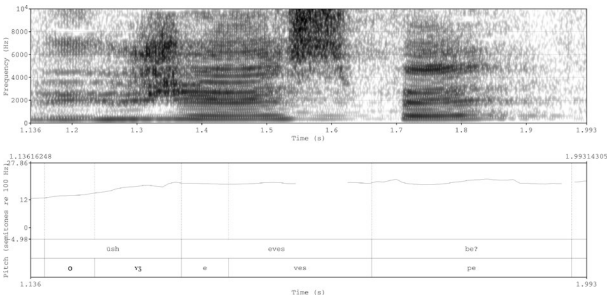


Figure 3. Intonogram of the utterance *Üsh eves be*? 'Not three times?'

Conclusions

The analysis of the Tuvan interrogative utterances allows us to make some preliminary conclusions:

1) The Tuvan yes/no questions are formed by adding the special interrogative particles or just by using a special intonation.
2) Unlike the obligatory rise of F0 on the interrogative particle **ba** in the Altay language (see Ryzhikova et al. 2020:149-165) in Tuvan the tone on **be** appears to be very inconsistent: it can be pronounced either with inclination, or even high tone or declination. The later phenomenon might be possibly explained by the position of focus, but this supposition requires further investigation.
3) The interrogative particle **be** in Tuvan is characterized by its lengthening, the phenomenon not found in Altay, where its duration equals the average syllable length (see Ryzhikova et al. 2020:149-165).
4) Although the beginning of the utterance can be the same for the question and the statement, then their intonation contours diverge: the question is characterized by inclination whereas a statement – by declination of F0.

All the conclusions are preliminary and need further verification.

Acknowledgements

This research is carried out within the framework of the project of the Institute of Philology of the Siberian Branch of the Russian Academy of Sciences «Cultural universals of verbal traditions of the peoples of Siberia and the Far East: folklore, literature, language» supported by a grant from the Government of the Russian Federation for the promotion of research conducted under the guidance of leading scientists, contract № 075-15-2019-1884.

References

Bavuu-Syuryun M. 2010. Tuvan language on modern stage: educational aspect. Novye Issledovaniya Tuvy 3, 57-71.

Johanson, Lars. 1998. The History of Turkic. In Lars Johanson & Éva Ágnes Csató (eds) The Turkic Languages. London, New York: Routledge, 81-125.

Scherbak, A. 1994. Introduction into a comparative study of the Turkic languages. Saint-Petersburg.

Simchit, A. K.-M. 2017. The Tuvan language in Big Russian Encyclopaedia; online: https://bigenc.ru/linguistics/text/4218682 Visited: 25.08.2020.

Sat, Sh. 1983. Modern Tuvan literary language. A textbook. Kyzyl. 141 p.

Ryzhikova, T., Dobrinina, A., Plotnikov, I., Shestera, E., Shamrin, A. 2020. Intonation of modal interrogative sentences in the Altai language (on the Altai-Kizhi folklore material). Siberian Journal of Philology 2, 149-165.

A data-driven caption for L2 listening

Maryam Sadat Mirzaei, Kourosh Meshgi
RIKEN Center for Advanced Intelligence Project (AIP), Japan
https://doi.org/10.36505/ExLing-2020/11/0042/000457

Abstract

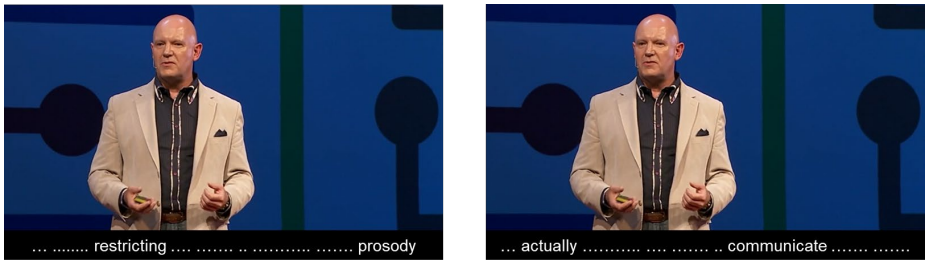
Partial and Synchronized Caption (PSC) is a tool that automatically detects difficult segments for the second language (L2) listeners and displays them in the caption while omitting easy-to-recognize cases to reduce cognitive load. Given that the number of words to be shown in this caption is limited, the main challenge lies in selecting and prioritizing difficult words. Since partialization is a classifying task, we made a dataset of labeled words in TED talks (easy vs. difficult) for a target proficiency-level. A deep classifier is trained on this dataset to automate the detection of difficult words/phrases without explicitly extracting word features. This proposed data-driven PSC outperforms its feature-based versions by adopting a selection pattern that is more similar to the annotations, capturing more complicated cases, and minimizing the false positives.

Keywords: listening difficulty, partial and synchronized caption, data-driven word selection

Introduction

The use of captions in training listening skills of L2 learners has been repeatedly studied. While full captions are considered helpful due to the simultaneous use of verbal and visual inputs according to dual coding theory (Paivio, 1990), it is criticized for overloading learner's cognitive load (Mayer & Moreno, 2003), split attention (Chang, 2009) and encouraging learners to read the text rather than listening to the audio, since it is easier for them (Leveridge & Yang, 2013).

To alleviate such problems, partial captioning has been introduced as a substitute for the full captions. The objectives of using partial captions in the literature are different. Keyword Caption (Guillory, 1998) presents only the keywords to the learner to facilitate their comprehension of the video, presenting verbal scaffold to learners regardless of their proficiency, listening difficulties, and the listening material. To address such shortcomings, we proposed Partial and Synchronized Caption (Mirzaei et al., 2018), which aims to foster L2 listening by promoting listening over reading and choosing the potentially problematic parts of listening material to appear in the caption. PSC is a type of captioning that synchronizes the words' appearance with the onset of their utterance while showing only the difficult words for the learners. This caption provides the minimal but necessary scaffold for L2 listeners to facilitate their listening, lower their reliance on reading, and promote the word boundary detection.



Original sentence: "*it's actually restricting your ability to communicate through prosody*"
How to speak so that people want to listen -Julian Treasure, TED 2013

Figure 1. Word selection for PSC: (left) shown words are labeled difficult for target learners, (right) the selected words are considered easy (and not useful) for them.

One of the challenges of making such a caption is to detect potential listening difficulties of the learners (Figure 1). Early versions of PSC use hand-crafted criteria to automatically omit easy-to-recognize words while identifying and showing difficult words to ease the cognitive load of L2 listeners. In those versions, difficult words/phrases were selected if they had a low frequency, high speech rate, or if they were considered as specific terminologies. Moreover, to add acoustically difficult words, it included special cases where automatic speech recognition errors and L2 learners' mistakes were highly correlated (e.g., breached boundaries). However, different aspects of a word in caption may contribute to its difficulty. Thus making an exhaustive set of rules to capture such complex relation is deemed intractable and daunting. By reframing the problem to an easy-difficult word classification, the problem can be solved with state-of-the-art machine learning tools.

To this end, we prepared an annotated TED speech corpus for a target L2 proficiency level, trained a deep classifier on the speeches' transcripts to classify its easy and difficult words, and generated PSC using this classification. This trained predictor is then used on unseen videos, and the word selection is examined against the rule-based PSC and the annotated labels (ground truth).

Method

We prepared a corpus of TED talk videos delivered by native speakers with a cumulative length of 93 minutes, including more than 10,000 words. The videos are forced aligned with Kaldi ASR system in word-level, and their show/hide labels are provided by two annotators targeting intermediate proficiency. Our annotators were L2 instructors with clear annotation guidelines and identical instructions on the criteria to label the words. The resulting annotation has κ=0.83 Cohen inter-annotation agreement.

These labels, along with the words, were used to train a classifier to predict difficult words for the target proficiency level of L2 learners. A deep text

classifier is trained to classify each word into easy and difficult categories. In this classifier, words are fed to an embedding layer that maps the vocabulary of the corpus into a compact representation, increasing the classifier's accuracy. To (partially) account for acoustic aspects of the difficulties, the word embedding is augmented with the speech rate feature. This augmented encoding is then fed to two stacked LSTM layers (Hochreiter & Schmidhuber, 1997). These layers capture the context in which the word is used and provide the encoded sequence context to the following fully connected layers. The two fully connected layers employ Batch Normalization and DropOut (0.5). The processed word is finally passed to a SoftMax layer, which decides if the word should be shown (i.e., difficult word for target proficiency) or not. The model is implemented in TensorFlow.

Results and discussion

We compared our proposed data-driven PSC with the feature-based methods. In the case of feature-based PSC, the pool of features covers lexical (e.g., frequency, specificity, length, syllables), syntactic (e.g., part-of-speech, dependency parse relations), semantic (e.g., polysomic words, co-references, idiomaticity), and acoustic or perceptual complexities (e.g., speech rate, breached boundaries, negatives). Three different classifiers (SVM, Naïve Bayes, and Decision Tree) have been trained on the dataset with all these features, and the labels annotators agreed upon. Each version of PSC predicts the difficulty labels, and their performances are compared with the rule-based PSC (with speech rate, specificity, and word frequency) via 5-fold cross-validation.

Table 1. Comparing the performance of proposed data-driven PSC with feature-based PSC on the annotated dataset for intermediate learners (%)

	Easy words		Difficult words		Metrics			
Caption	Show	Hide	Show	Hide	Precision	Recall	Accuracy	Sensitivity
PSC-Rule	11.5	56.6	15.3	16.6	57.1	48.0	71.9	83.1
PSC-DT	9.2	60.9	10.3	19.6	52.8	34.4	71.2	86.9
PSC-NB	**2.4**	**71.5**	12.5	13.6	**83.9**	47.9	84.0	**96.8**
PSC-SVM	8.5	69.5	15.1	6.9	64.0	68.8	84.6	89.1
Proposed	6.7	70.8	**16.3**	**6.7**	70.9	**72.4**	**87.1**	91.4

Table 1 shows that the Naïve Bayes classifier better handles the easy words, whereas the proposed method performs better in dealing with difficult words. It also shows that these two classifiers have superior performance compared to DT, SVM, and the rule-based PSC. Since the caption should prioritize showing difficult words to L2 learners, our proposed method is more desired.

By analyzing the errors of the classifiers, several interesting findings have been observed: *(i)* Some words are considered difficult in one context but not

others. Feature-based methods calculate lexical features (word frequency and specificity) regardless of the word context, which leads to some false positives in detecting difficult words, *(ii)* Since the annotators label difficult words based on the context, they consider some additional features (such as word surprisal that depends on the context) in labeling. The proposed method that uses the word context is capable of detecting such cases. *(iii)* Without adding a speech rate feature, the proposed method has difficulty in correctly identifying words that are acoustically difficult such as those uttered excessively fast. However, by adding the speech rate to the word encoding in our method, some of these cases are handled properly. On the other hand, the speech rate alone is not enough to cover a wide range of acoustically difficult words. Handling such cases requires the inclusion of more speech-driven features to the pipeline.

In summary, the proposed method is capable of predicting the listening difficulty of L2 learners for a given the word and is able to automatically generate PSC by observing a few instances of expert labels for a target L2 proficiency level. This method is extendable to allow for learner adaptation by providing annotations on different proficiency levels. It is anticipated that with more training samples, this data-driven PSC can provide a better form of assistance to foster listening to authentic materials for L2 learners at different levels.

References

Chang, A.C.S. 2009. Gains to L2 listeners from reading while listening vs. listening only in comprehending short stories. System, 37(4): 652–663.

Guillory, H.G. 1998. The effects of keyword captions to authentic French video on learner comprehension. Calico Journal, 15(1–3): 89–108.

Hochreiter, S., Schmidhuber, J. 1997. Long short-term memory. Neural computation, 9(8), 1735-1780.

Leveridge, A.N., Yang, J.C. 2013. Testing learner reliance on caption supports in second language listening comprehension multimedia environments. ReCALL, 25(2): 199–214.

Mayer, R.E., Moreno, R. 2003. Nine ways to reduce cognitive load in multimedia learning. Educational Psychologist, 38(1): 43–52.

Mirzaei, M.S., Meshgi, K., Kawahara, T. 2018. Exploiting automatic speech recognition errors to enhance partial and synchronized caption for facilitating second language listening. Computer Speech & Language, 49, 17-36, Elsevier.

Paivio, A. 1990. Mental representations: A dual coding approach. Oxford University Press.

Planning of active and passive voice in German

Judith Schlenter, Yulia Esaulova, Elyesa Seidel, Martina Penke
Department of Special Education and Rehabilitation, University of Cologne, Germany
https://doi.org/10.36505/ExLing-2020/11/0043/000458

Abstract

This eye-tracking experiment investigated how morphological case affects German speakers' descriptions of transitive events, specifically whether explicit case marking modulates speakers' structural choices. To increase the production of non-canonical structures (passive, patient-initial active), we primed patients in event scenes with a red dot. Subject and object case in German are unambiguously marked on masculine nouns but not on feminine nouns. If explicit case marking requires more structural planning, we should find an effect of gender. For feminine nouns, speakers may start with the cued patient and continue with a passive or a patient-initial active sentence. However, analyses of syntactic choice, speech onset times and eye gaze revealed that gender and thus case marking had no effect on sentence planning.

Keywords: sentence production, eye tracking, passive voice, structural choice

Introduction

Sentence-production studies combined with eye-tracking have demonstrated that directing speakers' attention to one of the characters in an agentive event scene by means of a perceptual cue can affect speakers' syntactic choice. In English, cueing of patients led to a significant increase in the number of passive voice utterances from 15% when the agent was cued to 26% when the patient was cued (Gleitman et al. 2007: Exp. 2). In languages with morphological case such as Russian, Finnish, or German, however, speakers seem considerably less inclined to produce a passive (percentage of passives 0-6%, e.g. Esaulova et al. 2019, Myachykov et al. 2011) – although eye-tracking analyses revealed that the cueing manipulation was effective and directed participants' attention towards the cued patient. Case-marking languages often offer the possibility to produce patient-initial active sentences such as OVS, a fact that might account for the disinclination to produce passive voice in these languages, especially in languages like Russian where passives are assumed to be strongly dispreferred (e.g., Tomlin & Myachykov 2015).

Our study investigated how morphological case affects the description of transitive events and the propensity to produce passives. The necessity to overtly case-mark an actant in an event as subject or object might affect sentence planning and thus account for the observed cross-linguistic differences with respect to structural choice. German is ideally suited for this investigation for two reasons. First, passives in German appear with a frequency comparable to English in language production corpora. Second,

subject (nominative) and object case (accusative) are distinctly marked on masculine nouns but not on feminine nouns (e.g., [der Vampir]$_{NOM}$ vs. [den Vampir]$_{ACC}$ 'the vampire' compared to [die Nonne]$_{NOM}$ vs. [die Nonne]$_{ACC}$ 'the nun'). We hypothesized that ambiguity in case marking may facilitate the application of a linear incremental planning strategy (Gleitman et al. 2007) and increase the production of non-canonical structures (passive or OVS). This is because participants could start utterance planning with the first increment, the cued feminine patient, still having the option to continue with an OVS active or an SVO passive sentence. In contrast, the necessity to overtly case-mark a masculine patient as subject or object of the sentence could lead to longer speech onset times or to fewer non-canonical structures.

Method

We tested 41 native speakers of German (10 male) in an eye-tracking study where they had to describe 24 black-and-white scenes depicting an agent and a patient character in a single sentence. We aimed to elicit non-canonical structures by perceptually priming the patient of the action by a red dot (presented for 700 milliseconds) that directed participants' attention to the position where the patient was to appear in the depicted scene. Additionally, we always presented animate patients to the left of animate agents as previous research has shown that position and animacy of the patient affect syntactic choice in German participants (Esaulova et al. 2019). We investigated the influence of case marking on syntactic choice by systematically varying the gender of agent and patient: in half of the items the scenes depicted two masculine role nouns, in the other half the actants were feminine role nouns. Masculine and feminine nouns were matched with respect to frequency and number of syllables. All nouns were simplex nouns. Agents and patients were depicted with similar size and visual complexity.

Twenty-four additional drawings displaying two figures or objects in different spatial configurations were presented to elicit different sentence structures (e.g. locative constructions) and constituted filler items.

Eye gaze was measured with an *EyeLink 1000 Plus* eye tracker at a sampling rate of 500 Hz. Participants' utterances were recorded with a headset with a boom microphone.

The data were analyzed with respect to (i) syntactic choice, (ii) speech onset time (SOT), and (iii) eye gaze. In total 936 trials were analyzed (exclusion of 4.9% of data due to unanalyzable utterances, disturbances, or self-corrections). Mixed-effects models were computed in R (R Core Team 2020). The best fitting model was selected based on the lowest AIC value.

Results

Participants almost exclusively produced SVO active sentences, only 7% of the utterances were passives. Participants did not produce any OVS active

sentences. There was no effect of gender on syntactic choice (β = 0.426, SE = 0.495, χ = 0.862, p = 0.389), nor was there an effect of gender on the reciprocal square root transformed SOTs (t = 0.005, p = 0.996).

Visual inspection of the eye gaze data indicated no difference between masculine and feminine nouns for active utterances, see Figure 1 (data on passives not displayed due to scarcity of data). The data also showed a higher likelihood of looks to the agent as compared to the patient after around 300 ms until shortly before speech onset. A mixed-effects logistic regression model analyzing fixations on patient and agent within 20 ms time bins of every active utterance trial between 200 and 600 ms showed that across gender conditions speakers were more likely to fixate the agent than the patient but there was no modulation by gender. The analysis window was chosen based on previous research on speech planning (e.g., Konopka et al. 2018).

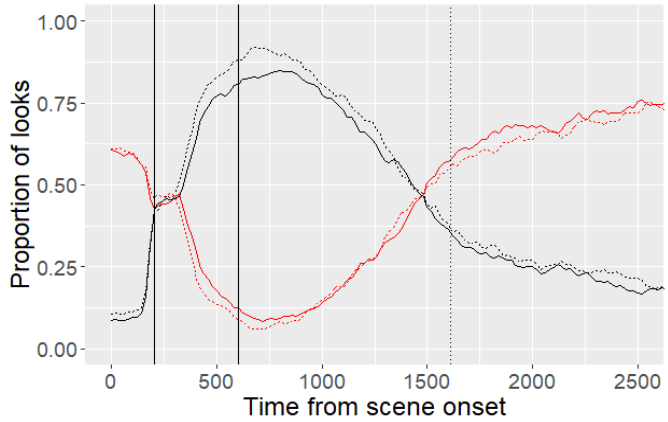


Figure 1. Proportions of fixations on agent (black) and patient (red) masculine (dashed lines) and feminine (solid lines) role nouns for active voice utterances. The dotted vertical line indicates the average speech onset, the two solid lines show the window chosen for statistical analysis.

Discussion

Although visual cueing was effective and directed our participants' gaze to the patient of an event scene, this manipulation of attention did not lead to a higher rate of non-canonical structures, confirming previous research that has found speakers of German to be disinclined to produce passives or OVS actives in event descriptions (Esaulova et al. 2019, 2020). The aim of our study was to find out if this cross-linguistic difference in the propensity to produce non-canonical structures was related to morphological case marking. However, our data provide no indication that explicit and unambiguous marking of morphological case on the sentence-initial NP affects syntactic choice in

German. We had assumed syncretic case marking on feminine NPs to lead to an increase of non-canonical sentence structures and faster SOTs, as participants could start with utterance production before the syntactic structure of the utterance is fully planned. In contrast, however, participants displayed a strong preference to focus on the agent and to realize it as subject of an active clause, even though their attention was directed to the patient.

Our results suggest that the observed cross-linguistic differences in passive production cannot be accounted for by explicit morphological case marking, by the presence of alternative syntactic structures such as OVS, or by a general avoidance of passive voice due to the infrequency of this structure.

Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 281511265 – SFB 1252. We are grateful to Sarah Dolscheid for helpful comments.

References

Esaulova, Y., Penke, M., Dolscheid, S. 2019. Describing events: Changes in eye movements and language production due to visual and conceptual properties of scenes. Frontiers in Psychology 10, 1–15.

Esaulova, Y., Penke, M., Dolscheid, S. 2020. Referent cueing, position, and animacy as accessibility factors in visually situated sentence production. Frontiers in Psychology 11, 1–13.

Gleitman, L., January, D., Nappa, R., Trueswell, J. 2007. On the give and take between event apprehension and utterance formulation. Journal of Memory and Language 57(4), 544–569.

Konopka, A. E., Meyer, A., Forest, T. A. 2018. Planning to speak in L1 and L2. Cognitive Psychology 102, 72–104.

Myachykov, A., Garrod, S., Scheepers, C. 2011. Perceptual priming of structural choice during English and Finnish sentence production. In Mishra R. K., Srinivasan, N. (eds.), Language-Cognition Interface, 53–71. Munich, Lincom.

R Core Team 2020. R: A language and environment for statistical learning (version 4.0.0). Retrieved from https://www.R-project.org/

Tomlin, R., Myachykov, A. 2015. Attention and salience. In Dabrowska, E., Divjak, D. (eds.), Handbook of Cognitive Linguistics, 31-52. Berlin, De Gruyter Mouton.

Individual differences in processing pseudo-inflected nonwords

Julia Schwarz, Mirjana Bozic, Brechtje Post
Department of Theoretical and Applied Linguistics, University of Cambridge, UK
https://doi.org/10.36505/ExLing-2020/11/0044/000459

Abstract

While the role of word stems has received much attention in morphological processing, the effects of inflectional suffixes on lexical access remain unclear. We address this gap as well as the contribution of individual differences on morphological segmentation with a visual priming experiment. Inflected and uninflected nonwords were preceded by a non-linguistic baseline string or the target's suffix/word-final letters (e.g. XXXXing → SMOYING). The results indicate that the suffix length is crucial for morphological effects to surface in visual priming and that morphological processing may be modulated by the individual's reading profile and vocabulary size. We interpret this as evidence for variable morphemic activation: morphological cues can facilitate visual access when rapid whole-word processing is unavailable. The theoretical implications are discussed.

Keywords: morphological priming, nonwords, individual differences

Introduction

Although the use of morphological information in visual word processing is well-documented (Amenta and Crepaldi 2012), the limited research on inflection priming effects (in contrast to stem and derivational priming) is inconclusive and raises questions about the mechanisms involved in accessing inflectional suffixes. In English, only two studies have investigated inflection priming, and only in the auditory domain: in contrast to Emmorey (1989), who did not find a significant morphological effect for the inflections *-ing*, *-ed* and *-es*, recently Goodwin Davies and Embick (2019) did find significant priming for plural nouns.

In this paper, we investigate the role of inflections through visual suffix priming in nonwords. Nonwords cannot be stored in the mental lexicon and as such allow us to reduce semantic interference from stems and instead focus on the inflections in question. In addition, we were interested to see whether morphological priming effects in nonwords are dependent on subjects' reading skill and the size of their existing lexicon, as measured by their response speed and error rate. Medeiros and Duñabeitia (2016) found that morphological effects in visually presented derived words were indeed modulated by individual reading speed: only slow readers, who have a lower level of orthographic skill, displayed a morphological priming effect. Similar effects could apply to

inflected words. Furthermore, vocabulary size and subjects' reading skills are likely to affect processing of novel information such as nonwords.

We hypothesized that a high error rate, which has been related to lower vocabulary knowledge (Yap, Balota, Sibley, Ratcliff 2012), as well as slower reading speed (Medeiros and Duñabeitia 2016) could lead to a stronger dependency on morphological information, given that automated whole-word processing necessitates rapid orthographic decoding. Breaking down unfamiliar items into more meaningful sub-lexical units might therefore facilitate reading for subjects with lower orthographic and vocabulary skills.

Methods

The online experiment was set up in Psytoolkit (Stoet 2010, 2017). Pseudo-inflected nonwords were preceded by suffix primes embedded in a non-linguistic string (e.g. XXXXing → SMOYING). Each trial consisted of a row of hashmarks, a 150ms prime, and a lexical decision task to the target word. We tested three English inflections, varying in orthographic length: the plural inflections *-s* and *-es*, and the verbal inflection *-ing*, each matched with uninflected controls (e.g. XXXXgle → SHANGLE). The 288 nonwords were created from the real word targets used in the lexical decision task and matched on the number of letters and syllables, phonological shape, number of orthographic neighbours, and bigram frequency, resulting in a well-balanced data set of phonotactically legal items.

Reaction times from 80 native English speakers were log-transformed and analyzed in R by fitting linear mixed effects models with subject and item random intercepts for each suffix type (vs. uninflected orthographic controls). P-values were calculated with the Satterthwaite approximation for degrees of freedom. Mean reaction times and error rates per participant were calculated based on the cleaned data set and then rescaled and centred.

Results

Our results indicate that inflected nonwords can be primed with suffixes presented in isolation. However, this depended on the type/length of the primes and participants' response patterns. The optimized models fitted individually for nouns with an *-s* suffix and nouns with an *-es* suffix did not show effects of priming or morphological complexity. Since items and controls were closely matched, and since there was also no significant priming or inhibition for orthographic controls, it is likely that the length of the prime is crucial for a reliable visual priming effect.

However, the orthographically longest prime, the *-ing* suffix, showed significant priming compared to simple controls of the same length (-3ms vs. +15ms). In addition to a significant interaction between priming and morphological complexity (b=.02, $t(6157)=2.15$, p=.031), the best model also included subjects' mean reaction time (b=.14, $t(73)=31.86$, p<.001) and a three-

way interaction between priming, morphological complexity, and subjects' error rates (b=.03, t(6164)=4.01, p<.001).

Following a median-split comparison, we further investigated the effects of error rates and response speed on morphological priming in the 3-letter overlap condition (*-ing* vs. simple targets). We found that only subjects with a high error rate, which could indicate a smaller lexicon size, showed morphological priming effects of suffixed targets (-13ms, b=.04, t(2859)=2.70, p=0.007; Figure 1, left). In addition, only slow readers, indicated by their mean response time, displayed morphological priming (-10ms, b=.03, t(3038)=2.18, p=.029; Figure 1, right). The inhibitory effect found for uninflected targets was unaffected by reaction times and error rates.

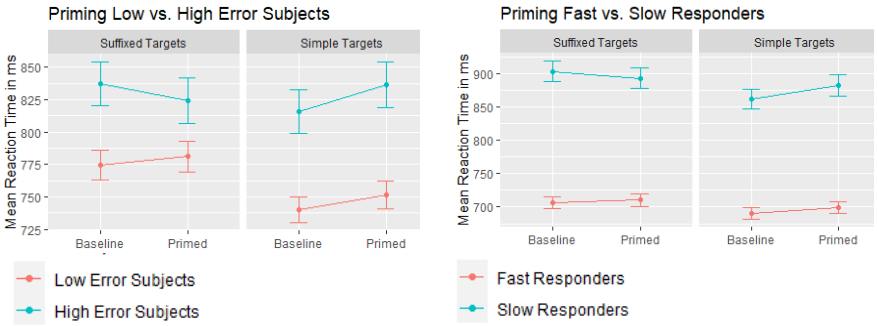


Figure 1. Priming effects for the *-ing* suffix, shown as a function of error rate and response speed.

Discussion

If inflectional suffixes play no role as sub-lexical units in visual access, no priming difference between inflections and orthographic controls should emerge, nor should there be any differences between subjects with different reading profiles and lexical knowledge. We found that the opposite was the case in our study, at least for the longest and therefore most salient inflection.

Specifically, the results suggest that the effects of morphological cues in the visual modality are dependent on both linguistic factors and individual variables. Firstly, nonwords were affected differently by a prime dependent on the affix structure; nonwords inflected with the *-ing* suffix were facilitated by a prime, whereas uninflected controls were inhibited. Secondly, morphological priming was modulated by subjects' response speed and error rates. Only subjects with high error rates and slow responses displayed morphological facilitation for suffixed targets. These results are in line with previous research which found that priming differences between morphologically derived and simple targets only emerged for readers with lower proficiency (Beyersmann, Casalis, Ziegler, Grainger 2015, Medeiros and Duñabeitia 2016). It seems that subjects with a smaller vocabulary (indicated by errors) and lower orthographic

processing skill (indicated by response speed) might be less efficient in mapping orthography onto meaning directly and as such are more likely to engage decomposition mechanisms.

Overall, the results indicate that morphemic chunks can facilitate lexical identification, especially for lower-skilled readers. We suggest that this reflects morphological processing as a sub-lexical strategy when direct mapping is unavailable. This finding is most consistent with parallel-route models, which allow competition between direct access and decomposition (e.g. Baayen and Schreuder 1999).

Acknowledgements

This work was supported by a Gates Cambridge Award and an ESRC dtp scholarship to the first author.

References

Amenta, S., Crepaldi, D. 2012. Morphological processing as we know it: An analytical review of morphological effects in visual word identification. Frontiers in Psychology 3, 232.

Baayen, H., Schreuder, R. 1999. War and peace: Morphemes and full forms in a noninteractive activation parallel dual-route model. Brain and Language 68(1-2), 27-32.

Beyersmann, E., Casalis, S., Ziegler, J., Grainger, J. 2015. Language proficiency and morpho-orthographic segmentation. In Psychonomic Bulletin & Review 22(4), 1054-1061.

Emmorey, K. 1989. Auditory morphological priming in the lexicon. Language and Cognitive Processes 4(2), 73-92.

Goodwin Davies, A., Embick, D. 2019. The representation of plural inflectional affixes in English: evidence from priming in an auditory lexical decision task. Language, Cognition and Neuroscience 60, 1-9.

Medeiros, J., Duñabeitia, J. 2016. Not everybody sees the ness in the darkness: Individual differences in masked suffix priming. Frontiers in Psychology 7, 1-10.

Stoet, G. 2010. PsyToolkit - A software package for programming psychological experiments using Linux. Behavior Research Methods 42(4), 1096-1104.

Stoet, G. 2017. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. Teaching of Psychology 44(1), 24-31.

Yap, M., Balota, D., Sibley, D., Ratcliff, R. 2012. Individual differences in visual word recognition: insights from the English Lexicon Project. In Journal of experimental psychology. Human Perception and Performance 38(1), 53-79.

Experimental testing of the left periphery

Jose Sequeros-Valle
Department of Hispanic and Italian Studies, University of Illinois at Chicago, US
https://doi.org/10.36505/ExLing-2020/11/0045/000460

Abstract

This manuscript presents an empirical description of the left periphery based on the performance of speakers of Castilian Spanish in a corpus analysis, an acceptability judgment task, and a scripted production task. The picture drawn by the three studies look as follows: First, clitic-doubled left dislocations (CLLD) fulfil multiple discourse functions, but the construction is not completely free from discourse restrictions. Second, canonical utterances are also able to fulfil CLLD's discourse functions. Third, CLLD does not present distinctive intonational patterns depending on the discourse function. Fourth, there is partial evidence that focus fronting (FF) presents an intonational pattern different than that of CLLD. The concluding section of the manuscript calls from a new model of the left periphery.

Keywords: clitic-doubled left dislocation, focus fronting, left periphery, information structure, Spanish

Background

Much generative research in recent years has focused on the interaction of different linguistic subdomains, and how this interaction can help us develop a holistic model of the architecture of human language: Does phonology have its own interpretation system or it depends directly on morphology? Are semantics and discourse two different domains or two aspects of the same domain? Is discourse pre- or post-syntactic? One subdomain extensively studied is discourse, on its interface with syntax and phonology. The interface among these three domains is worth investigating because it has implications for the larger questions presented above, and for language at large.

In order to look into the discourse-syntax-phonology interface, I take the Spanish left periphery as a test case; namely clitic-doubled left dislocations (CLLD: *A Pedro María lo ve en el parque*, 'Pedro, Maria see him in the park'), focus fronting (FF: *A Pedro ve María*, 'Pedro Maria sees'), and their canonical counterpart (María ve a Pedro, 'Maria sees Pedro). In general terms, FF expresses contrastive focus and presents an emphatic phonology (low intermediate boundary tone, with an extended pitch range), while CLLD expresses topicality and presents a clitic pronoun doubling the fronted constituent.

Despite this general description, previous accounts disagree on the details. First, Bianchi and Frascarelli (2010) argue that CLLD can fulfil multiple discourse contexts. Further, these discourse functions cannot be fulfilled by

canonical utterances, and each interpretation corresponds to a unique intonation. In a second approach, López (2009) argues for a unique interpretation of each FF and CLLD, and these two discourse functions can be fulfilled by canonical utterances as well. Third, Rubio Alcalá (2014) argues for a completely interpretational freedom of all canonical and non-canonical utterances. Further, none of these authors have provided experimental evidence for their claims.

Based on the disagreements presented above, I propose four research questions for experimental testing: **RQ1:** Does Spanish CLLD fulfil multiple discourse contexts? **RQ2:** Do canonical utterances fulfil the same discourse contexts as Spanish CLLD? **RQ3:** Are different discourse-contexts marked via specific-pitch accent in Spanish? **RQ4:** Is contrastive focus marked via contrastive-stress in Spanish?

Study 1: A corpus analysis

The findings discussed in this section come from the Spanish section of the NOCANDO Corpus (Brunetti et al., 2011; available via https://parles.upf.edu/llocs/nocando/home). The corpus presents 32 potential instances of CLLD and FF, but some of them did not include the characteristics described above. Therefore, I only present the findings from the remaining 20 instances of CLLD, and no instances of FF.

These 20 remaining instances of CLLD were coded following the guidelines from Götze et al. (2007), which showed to fulfil multiple discourse contexts (e.g., given, subset, etc.). Additionally, a qualitative analysis of the corpus indicated that canonical utterances also fulfil the same six discourse functions found for CLLD. The same 20 instances of CLLD presented multiple pitch-accents on an analysis in Praat (e.g., L+H*, L*+H, H+L*, etc.).

Study 2: An acceptability judgment task

Twenty-five monolingual speakers of Castilian Spanish participated in this study. The experimental stimuli consist of a 3x2 design: Three discourse conditions – local antecedent, new referent, and wh-answer - and two syntactic conditions - utterances with fronted versus non-fronted constituents.

A linear mixed model with the maximal random structure (by-participant slope and intercept, by-item slope and intercept) converged when conducted in R-Studio. The inferential results show significant effects by Discourse ($F(2,26) = 17.79$, $p < 0.001$) and Syntax ($F(1,33) = 7.14$, $p < 0.01$), as well as a significant Discourse*Syntax interaction ($F(2,18) = 3.45$, $p < 0.05$). Further, the Tuckey HSD pairwise comparisons show that CLLD is significantly dis-preferred in the wh condition in comparison to the local ($p < 0.001$, $g = 1.11$) and new ($p < 0.02$, $g = 0.67$) conditions, with a non-significant different between these two ($p = 0.30$, $g = 0.31$). Further, there are no significant differences between CLLD and canonical utterances in the two conditions

where CLLD is preferred; namely, the local ($p = 0.97$, $g = 0.14$) and new ($p = 0.13$, $g = 0.57$) conditions.

Study 3: A scripted production task

Twenty-two monolingual speakers of Castilian Spanish participated in this study; twelve in Session 1 and ten in Session 2. The experimental stimuli consist of a 3x2 design: Three discourse conditions – contrastive-focus (CF), contrastive-topic (CT), given-topic (GT) - and two syntactic conditions - fronted constituents in Session 1 versus non-fronted constituents in Session 2.

Table 1 presents the results from R-Studio, indicating the session, the variable, the analysis, the converging random structure, the intercept effect, and the three pairwise comparisons for each of the analyses.

Table 1: Inferential results of the scripted production task

Session	Variable	Analysis	Random Structure	Sig.?	CF-CT	CF-GT	CT-GT
1	Accent	GLMM	Intercepts	*	n.s.	n.s.	n.s.
1	Tone	GLMM	Intercepts	*	*	*	n.s.
1	Range	LMM	Maximal	*	n.s.	*	n.s.
2	Accent	GLMM	Intercepts	n.s.	n.s.	n.s.	n.s.
2	Tone	GLMM	Intercepts	*	*	*	*
2	Range	LMM	Maximal	*	n.s.	n.s.	n.s.

Discussion

I argue that the corpus results show that CLLD receives multiple discursive interpretations (RQ1). Second, the findings from the qualitative comparisons between CLLD and its canonical counterpart, further indicate that canonical utterances do fulfil the same discourse context as CLLD utterances do (RQ2). Third, the discourse-phonology results show that there is not a one-to-one correlation between the discourse function and the intonation of a fronted constituent (RQ3). Given the lack of actual instances of FF in the corpus, there is no evidence for or against the presence of a contrastive intonation for FF (RQ4).

For the acceptability results, first, the significant difference between wh-CLLD and the other two CLLD-types indicates that, while CLLD is acceptable in multiple discourse contexts (local antecedent and new referent), the interpretation of CLLD is not completely free (wh-answer) (RQ1). Further, the lack of significant differences between CLLD and its canonical counterparts in the local and new conditions indicate that canonical utterances do in fact fulfil the same discourse functions as CLLD does (RQ2). These findings provide further confirmation of the discourse-syntax findings from Study 1.

For the production results, first, the lack of significant differences in the pairwise comparisons show that different discourse functions do not correlate with different pitch accents (RQ3); as in Study 1, Study 3 did not find a one-to-one correlation between discourse and pitch accent. Second, contrastive-focus is significantly different than both topic-types for boundary tones in Session 1, which aligns with a positive answer to RQ4. The results for boundary in Session 2 and for range in both sessions do not indicate a clear pattern for a positive nor a negative answer to RQ4.

These findings have extensive implications. First, the flexibility in word order found in Studies 1 and 2, as well as the lack of correlation between discourse and pitch-accent in Studies 1 and 3, challenge Bianchi and Frascarelli's (2010) empirical predictions. Second, the fact CLLD is not accepted as an answer to a wh-word shows that this construction is not completely free from discourse restrictions, unlike predicted by Rubio Alcalá's (2014). Lastly, the fact that CLLD is accepted both when referring to a local antecedent as well as a to a newly introduced referent, challenges López's (2009) predictions.

Putting it all together, none of the three previous models on the discourse-syntax-phonology interface truly capture the empirical picture resulting from the three studies presented in this manuscript. Therefore, a new model of this 3-way interface is necessary. This task, however, is left for future inquiry (e.g., Sequeros-Valle, forthcoming) given the descriptive nature of the present paper.

References

Bianchi, V., Frascarelli, M. 2010. Is topic a root phenomenon? Iberia: An International Journal of Theoretical Linguistics, 2(1), 43-88.

Brunetti, L., Bott, S., Costa, J., Vallduví E. 2009. NOCANDO: A multilingual annotated corpus for the study of Information Structure. Presentation at the 5th Corpus Linguistics Conference. The University of Liverpool.

Götze, M., Weskott, T., Endriss, C., Fiedler, I., Hinterwimmer, S., Petrova, S., Schwarz, A., Skopeteas, S., Stoel, R. 2007. Information structure. In S. Dipper, M. Götze, S. Skopeteas (Eds.), Information Structure in Cross-Linguistic Corpora (pp. 147–187). Potsdam: Interdisciplinary Studies of Information Structure.

López, L. 2009. A derivational syntax for information structure. Oxford, England: Oxford University Press. · Rubio Alcalá, C. (2014). Syntactic constraints of topicalization phenomena. Ph.D. dissertation, Universitat Autònoma de Barcelona.

Sequeros-Valle, J. (forthcoming). [+a], [+p], and [+c]: A typology of discourse in Spanish. Typology of Morphosyntactic Parameters.

Prosody of focus in statements of the Altai language

Elena Shestera, Nikolay Urtegeshev, Iraida Selutina, Anton Shamrin
Laboratory of Phonetics, Siberian Branch of Russian Academy of Science, Russia
https://doi.org/10.36505/ExLing-2020/11/0046/000461

Abstract

The prosody of verbal word forms in the narrative utterances of the Altai language is under consideration in the article. In this work, in addition to the acoustic analysis in the Praat program, we took into account the subjective perception of native speakers. In the simple statements the intonation declines on the predicate when realizing the topic of the utterance. The focus of the utterance may be expressed by pitch and intensity peak.

Keywords: Altai language, intonation, prosody, narrative statement, topic, focus

Introduction

Intonation serves as a means of forming the communicative type of a sentence and syntactic relations between parts of a complex sentence and performs a delimitative and text-forming function. Since the expression of syntactic relations and the communicative type of an utterance is one of the functions of intonation, let us consider the relationship between intonation and information structure expressed in the prosody of verbal word forms of the Altai language.

The information structure of an utterance is expressed by the topic and focus. The topic corresponds to the logical subject of the proposition, while the focus is rather new or contrastive information that was stated or asked about. The focus contains new information and has the highest degree of communicative dynamism, but the completeness of information, or givenness, is created by a dynamic combination of the focus and topic (Dufter, Gabriel 2016). We could assume that the intonation contour of the focus differs from that of the topic, speech production being dependent on the sentence information structure is quite predictable, and it may be characterized as 'the topic intonation' or 'the focus intonation'.

An Altai utterance having the fixed word order S-O-P, in contrast to the Russian one, suggests the presence of certain intonation for expressing the topic and different intonation for the focus of the utterance. The material under investigation is a simple two-part (and simple one in complex sentences) statement with the topic and focus, where the subject is topic, and the predicate or object is focus. For each sentence (depending on the context), it is decided separately whether the predicate (which is expressed by a verbal word form and is a compound or nominal verb) represents the topic or focus of the utterance.

Material and methods

The research material was audio recordings of folk works of Altaians, i.e. Altai-Kizhi (Monuments of folklore... 2002, 2011). Audio recordings of the Altai folk tale and non-folktale prose fragments produced by four native speakers were listened to. Then utterances were selected for analysis. Our preference was given to simple uncomplicated sentences, but complicated sentences with homogeneous terms, and complex sentences were also under consideration.

Using an audio converter, all audio files were converted from WMA to WAV format. In the Audacity program, the sound signals were cut into separate utterances of the declarative communicative type. The resulting working files were annotated in the Praat program at several levels, but the material – 100 simple statements – was analyzed only at the syllabic sublevel.

Results

Consider some sample phrases pronounced by the native Altaian speakers.

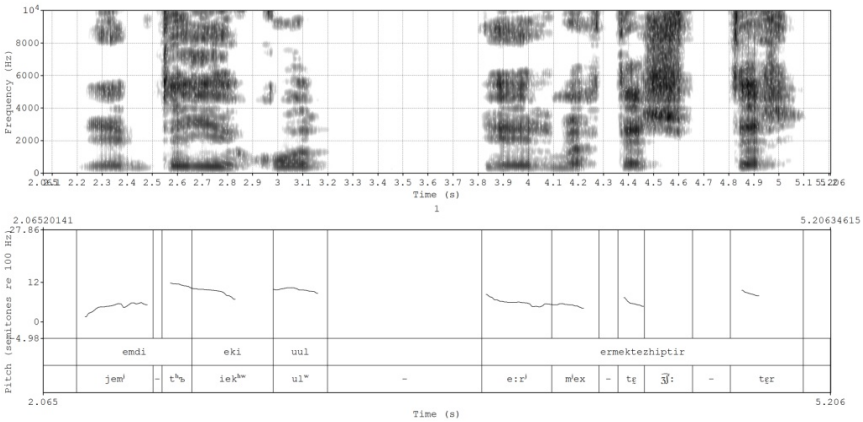


Figure 1. *Emdi eki uul ermektezhiptir*. – 'Here two brothers are speaking.'

The sentence in Figure 1 is simple, uncomplicated, between the group of the subject (which includes the adverbial modifier, the subject and the attribute) and the predicate we find a physical pause. The first intonation phrase – the group of the subject *emdi eki uul* – has an ascending-descending pitch. The maximum of F0 is on the second syllable of the adverbial *emdi* 'here'. On the verb, the pitch is generally lowered. An unexpectedly high F0 value appears on the final syllable of the predicate – perhaps it occurs as a result of drawing the listener's attention to what the narration will be about further, i.e., the predicate is the focus in the sentence.

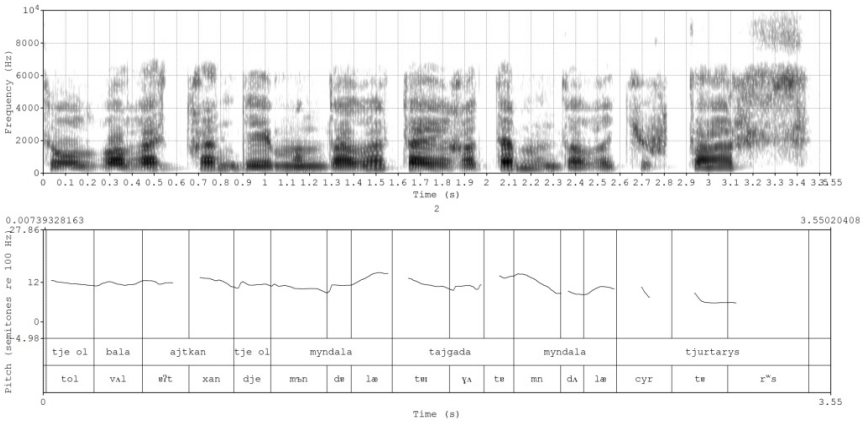


Figure 2. *Tje ol bala ajtkan, tje ol myndala tajgada, myndala tjurtarys*. – 'But this girl said, but here in taiga, here will we live.'

The sentence in Figure 2 begins with the conjunction *tje* 'but'. In the sentence, there is a lexical repetition of the adverbial *myndala* 'here'. The intonation contour is not stable: there are two F0 maxima occurred in both cases of the adverbial *myndala* 'here'. In the first case, the pitch increases, then decreases; on the repeated word *myndala*, the pitch decreases from high F0 value to low one in the utterance, reaching the lowest F0 value on the predicate in the final position of the sentence. Thus, the focus can be expressed not only by nuclear accent, or raising pitch, but also by pitch movements, i.e., from relatively high to minimal ones. This finding is consistent with the idea of R. Ladd that pitch increases are functional and may convey focus, and they are significant not only as individual components of intonation (out of context), but as a part of the global 'rise-fall' contour of an utterance (Ladd 1995: 117).

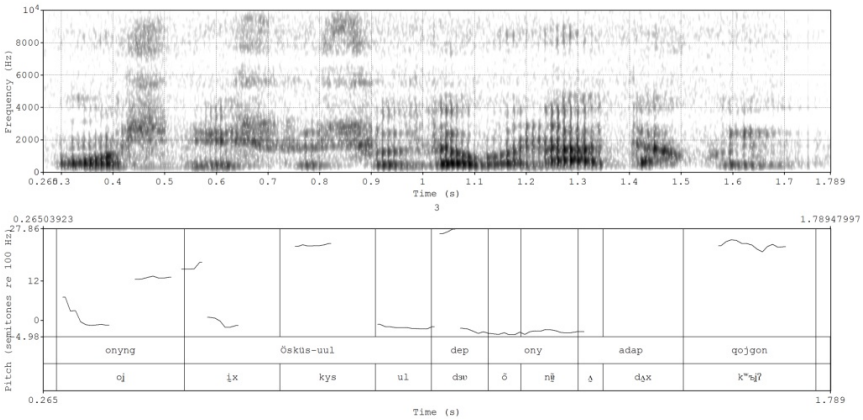


Figure 3. Onyng uchun Ösküs-uul dep ony adap qojgon. – 'Thus Ösküüs-lad was named like that'.

Here in this sentence in Figure 3, the compound conjunction *onyng uchun* 'thus' is found, and together with the verbal word *dep* 'having said' they may be attributed to the metatextual linkers. Being emphasized with the maximum pitch value, these contextual words realize the so called contrastive focus of this utterance; for further information the reader is referred to (Dufter, Gabriel 2016: 6). On the verb *ada*= 'to give name' we can observe declining F0 contour, which has the minimum pitch value at the end of the utterance. It is interesting that not only the F0 maximum but intensity peak is realized on the verbal word *dep* 'having said' as well. This metatextual linker originating from participial form of the verb de= 'to say' in folk texts often performs the function of strengthening declarativeness (of the statements) (Tybykova 1991: 175).

Conclusion

Thus, in the Altai language, the rise-fall intonation contour is mainly observed in declarative utterances. In a simple statement characterized by the standard for Altai word order S-O-P, the intonation declines on the predicate when realizing the topic of the utterance. The focus of the utterance may be expressed by pitch and\or intensity peak, or pitch movement from high value to lower one on the syllable of the emphasized word of the utterance.

Acknowledgements

This research is carried out within the framework of the project of the Institute of Philology of the Siberian Branch of the Russian Academy of Sciences «Cultural universals of verbal traditions of the peoples of Siberia and the Far East: folklore, literature, language» supported by a grant from the Government of the Russian Federation for the promotion of research conducted under the guidance of leading scientists, contract № 075-15-2019-1884.

References

Dufter A., Gabriel Ch. 2016. Information structure, prosody, and word order. In Fischer, S., Gabriel, Ch. (Eds.), Manual of Grammatical Interfaces in Romance 10. Berlin: De Gruyter.

Ladd R. 1995. 'Linear' and 'Overlay' descriptions: An autosegmental-metrical middle way. Proceedings of XIIIth International Congress of Phonetic Sciences, 116-123. Stockholm, Sweden.

Monuments of Folklore of the Peoples of Siberia and the Far East. Vol. 21: The Altai folktales. Novosibirsk, Nauka, 2002, 455 p. (in Russian).

Monuments of Folklore of the Peoples of Siberia and the Far East. Vol. 30: The Altai non-folktale prose. Novosibirsk, Nauka, 2011, 576 p.(in Russian).

Tybykova A. T. Research on the Syntax of the Altai Language: a Simple Sentence. Novosibirsk, 1991, 174-178. (in Russian).

Zipf's law in Toki Pona

Dariusz Jan Skotarek
Institute of Applied Linguistics, University of Warsaw, Poland
https://doi.org/10.36505/ExLing-2020/11/0047/000462

Abstract

Zipf's Law states that within a given text the frequency of any word is inversely proportional to its rank in the frequency table of the words used in that text. It is a statistical regularity of a power law that occurs ubiquitously in language – so far every language that has been tested was found to display the Zipfian distribution. Toki Pona is an experimental artificial language spoken by hundreds of users. It is extremely minimalistic – its vocabulary consists of mere 120 words. A comparative statistical analysis of two parallel texts in French and Toki Pona showed that even a language of such scarce vocabulary adheres to Zipf's Law just like natural languages.

Keywords: Zipf's Law, Toki Pona, artificial languages, computational linguistics, statistics

Introduction

It is well-known that language users tend to choose certain words more often than others. George Kinsley Zipf discovered that it is a regular tendency, which can be accurately predicted with a mathematical formula. Zipf observed that having ranked the words in a text according to the number of their occurrences, different frequencies of ranked words create a harmonic series 1, 1/2, 1/3, 1/n. In other words, in a body of text of substantial volume the second most-frequent word will appear half of the times as the first most-frequent, third one will appear one third as often, etc. This means that the amount of times a word is used is proportional to 1/rank. This statistical method of language description turned out to be ubiquitous and accurate for all languages that have been tested so far – from English and Chinese to Esperanto and Meroitic. Zipfian regularity plotted on a logarithmic scale demonstrates a power-law distribution, as showed in Figure 1.

Toki Pona is an experimental philosophical artificial language created by Canadian linguist Sonja Lang in 2001. It is known amongst other constructed languages for its very limited vocabulary – its lexicon consists of only around 120 words. It is an extremely simple language. There are very few grammatical rules and no irregularities. In the light of its limited vocabulary, word formation takes place via combining elements, such as jan pona 'good person' meaning 'a friend'. It is a living language – there are hundreds of fluent Toki Pona speakers. It is gaining popularity – since the academic year 2020/2021 the University of Geneva will be offering a Toki Pona course.

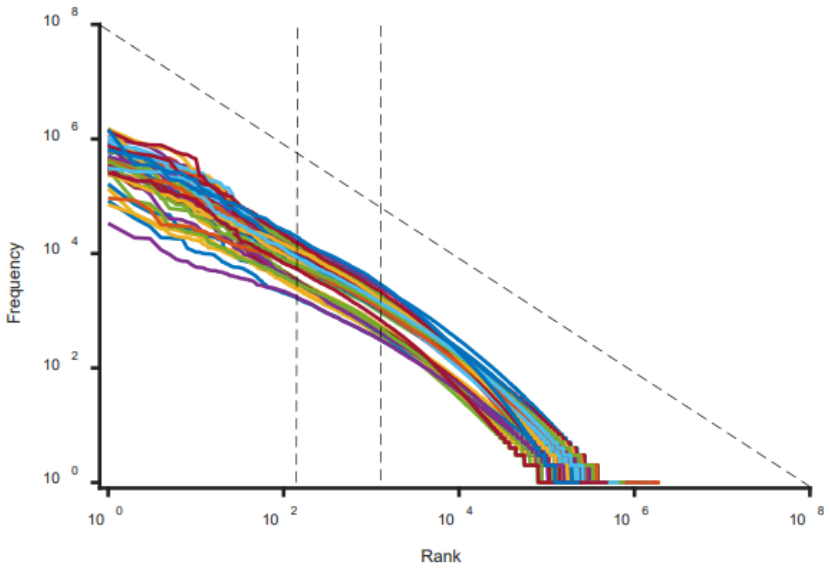


Figure 1. Frequency-rank distribution of 50 languages from various language families.

It is clear that Toki Pona differs tremendously from natural languages as well as from other artificial ones. Its vocabulary range of 120 words is nothing compared to that of, for example, English (approx. 250 000). If Zipf's Law is observed in languages with a wide vocabulary range, one could assume that Toki Pona could possibly "disobey" Zipf's Law. A natural language text consists of a handful of words that are used very often (i.e. articles and prepositions in English) but also of numerous hapax legomena – words used just once. This tendency makes natural languages "Zipfian". However, there are very few hapax legomena in Toki Pona – numerous words are used very often, since a given lexical unit constitutes a part of various lemmas. The point of the study was to "confront" Zipf's Law with Toki Pona.

Methodology

Two parallel texts were analysed – a legend "The life of Merlin" by Robert de Boron in the original, French version and its translation into Toki Pona done by a fluent Toki Pona speaker. Each text was parsed into a string of separate words. However, one aspect of the French language presented a problem in word-frequency analysis, namely elision, which is also marked in writing, for example l'ami 'friend' standing for le + ami. It raises a question whether one should view such cases as one or two words; an issue critical for a valid frequency-distribution analysis. In line with the Zipfian perception of what

constitutes a separate word, such instances were separated into two units. Parsed words were ranked according to the number of occurrences. Hence the rank-frequency distributions of the two texts were established.

Subsequently, those actual distributions were compared to those texts' perfect distributions predicted by Zipf's Law. The resulting correlation between the two values is a measure showing to what extent a given text obeys Zipf's Law.

Results

The French text's distribution represented a correlation with the Zipfian prediction of =0.89, with =1 being a perfect match. That result was in line with previous statistical analyses of natural languages. The Toki Pona text – contrary to possible initial assumptions – did obey Zipf's Law as well. Moreover, its distribution was coherent with the Zipfian prediction to the degree of =0.95, which is significantly higher than that of the French text.

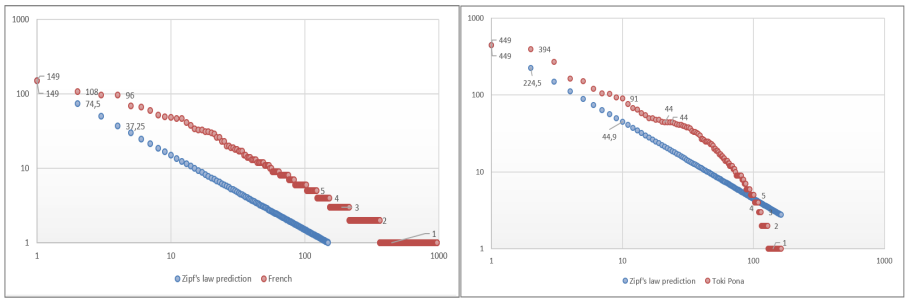


Figure 2. French (left) and Toki Pona (right) texts' lexical distributions and their respective Zipfian predictions.

Conclusions

The results of the study confirmed Zipf's Law's universality across both natural and artificial languages – even such unique, experimental ones like Toki Pona. It can constitute a valid argument in the discussion on possible explanations of Zipf's Law. One of such explanations is that every language is Zipfian because of its speakers who change it over time with their usage, which is guided by the principle of least effort. However, Toki Pona is a relatively young language, which did not undergo such user-driven change. Another explanation is that a Zipfian distribution is a reflection of a cognitive pattern that shapes human thinking – so that even while creating an artificial language its author unconsciously adhered to Zipf's Law. This thesis is supported by the fact that Sonja Lang confirmed that she was not aware that such law existed. Zipf's Law manifestations in phenomena other than language – such as city populations, web traffic or even size of Pluto craters – is a reason to assume that Zipf's Law is a purely statistical occurrence, with no deeper reason behind it. However, this

fact can also be used to pose a contrary argument – that Zipf's Law is a sign of something bigger, which eludes human understanding. Nevertheless, linguists shall test subsequent languages to see if indeed all human languages obey Zipf's Law, since – given the sheer quantity of languages on Earth – there are still many languages that remain untested in this respect.

References

Jiang, B., Yin, J., Liu, Q. 2015. Zipf's Law for All the Natural Cities around the World. International Journal of Geographical Information Science 29, 1–20.

Lang, S. 2014. Toki Pona: the language of good – the simple way of life. United States, Sonja Lang.

Reed, W.J., Hughes, B. D. 2002. From gene families and genera to incomes and internet file sizes: Why power laws are so common in nature. Physical Review E 66, 1–4.

Scholkmann, F. 2016. Power-Law Scaling of the Impact Crater Size-Frequency Distribution on Pluto. Progress in Physics 12(1), 26–29.

Smith, R. 2007. Investigation of Zipf-plot on the extinct Meroitic language. Glottometrics 15, 53–61.

Yu, S., Xu, C. Liu, H. 2018. Zipf's law in 50 languages: its structural pattern, linguistic interpretation, and cognitive motivation. Semanticscholar.org.

Zipf, G.K. 1935. The Psycho-Biology of Language: An Introduction to Dynamic Philology. Cambridge, The MIT Press.

Zipf, G.K. 1949. Human Behavior and The Principle of Least Effort. Cambridge, Addison-Wesley Press.

Academic vocabulary in an English for Academic Purposes course

Sophia Skoufaki[1], Bojana Petrić[2]
[1]Department of Language and Linguistics, University of Essex, UK
[2]Department of Applied Linguistics and Communication, Birkbeck, University of London, UK
https://doi.org/10.36505/ExLing-2020/11/0048/000463

Abstract

Academic vocabulary instruction can be beneficial to students in EMI universities since academic vocabulary knowledge predicts performance in academic tasks. With the aim to inform EAP materials design, this study examines the occurrence and repetition of high-frequency academic vocabulary in the printed teaching materials used in a presessional EAP course at a UK university. Findings indicate that even when EAP teachers do not design materials with the intention to include high-frequency AVL lemmas, as indicated from the interviews, they do include many. However, the average repetition rate of academic vocabulary was below 10 occurrences and, hence, unlikely to lead to incidental vocabulary learning. Implications for research and pedagogy are discussed.

Keywords: EAP, teaching materials, vocabulary, vocabulary repetition

Introduction

Most studies on the effect of English for Academic Purposes (EAP) courses on vocabulary learning and use compare students' pre- and post-course vocabulary knowledge (e.g., Alothman, 2014) or the lexical qualities (e.g., accuracy, variation) of their writing (e.g., Mazgutova & Kormos, 2015). Since these studies did not examine academic vocabulary instruction, it is unclear whether the improved vocabulary use or knowledge found between pre-tests and post-tests is due to vocabulary instruction in an EAP course or to other factors such as students' own independent vocabulary learning or their immersion in an English-speaking environment.

This study[1] takes the first step towards filling this research gap. It examines the occurrence and repetition of high-frequency English academic words, that is, words occurring in academic writing and speech across disciplines, in the teaching materials used in a presessional EAP course at a British university.

Research questions

Given the central role teaching materials play in language learning and the very limited research on the academic vocabulary students are exposed to through teaching materials used in EAP courses, this study aims to examine the

academic vocabulary in the materials used in an EAP course guided by the following questions:

1. How many high-frequency academic word lemmas appear in the EAP materials used in a UK presessional university EAP course?
2. The study focuses on vocabulary for academic reading and writing purposes. Recent research indicates that recall (not recognition) knowledge of vocabulary best predicts performance not only in writing but also in reading. Given that at least 10 occurrences of a word are necessary for recall knowledge to be developed via incidental vocabulary learning (Uchihara et al., 2019), how many high-frequency academic word lemmas appear in EAP materials at least 10 times?

Method

High-frequency academic word lemmas are operationalized as the Academic Vocabulary List (Gardner & Davies, 2014), which lists the most frequent 3,014 academic lemmas in the Corpus of Contemporary American English (COCA). The AVL consists of words from all frequency bands of COCA which occur at least 50% more frequently in the Academic section of COCA than would normally be expected, are evenly distributed across its disciplinary sections and occur in at least seven of COCA's eight disciplinary sections. The first AVL band and the second AVL band each consist of 1,000 lemmas and 'the third AVL band' consists of 1,014 lemmas.

This study identified AVL lemmas in the materials used in a presessional course at a British university. Teaching was structured in terms of the language skills and language knowledge that students need to develop to perform well in tasks at university. Three modules were taught, each by a different teacher: Reading and Writing (Reading/Writing), Listening and Speaking (Listening/Speaking), and Vocabulary and Grammar (Vocabulary/Grammar).

At the end of each week for five weeks, each EAP teacher provided us with photocopies of the teaching materials used by him/her that week and was interviewed about the content taught that week and how the teaching materials were used. Files were pre-processed, tagged for part of speech and searched for AVL lemmas.

Results and discussion

The AVL coverage of lexical word (i.e., noun, verb, adjective and adverb) tokens was examined because the AVL contains only lexical words. AVL lemmas from the first frequency band provided nearly all of the AVL coverage of the materials used in every module.

A total of 846 AVL lemmas were identified in the materials. Most of them (675) appeared in the written materials of the 'Reading/Writing' and 'Vocabulary/Grammar' modules and 199 were shared between these modules.

In relation to the second research question, a distinction was made between words which occurred in the materials 10 or more times and those which occurred less than 10 times because we aimed to see whether the materials provided students with enough lemma occurrences so that they were likely to be able to recall the meaning and form of at least one third of the academic lemmas in the materials. Second, a distinction was made between words which occur only once and those which occur more than once because the former are less likely to be learned than the latter (e.g., Waring & Takaki, 2003). Consequently, Band 1 included the AVL lemmas occurring once, Band 2 included those occurring 2-9 times and Band 3 included those occurring 10 or more times.

AVL lemmas that occurred 2-9 times form the majority (45.59%) of AVL lemmas in the materials. Only 14.42% of all the AVL lemmas in the materials occur 10 or more times; all but three come from the most frequent 1,000 lemmas in the AVL. Therefore, an AVL lemma was more likely to appear 10 or more times in the materials if it was among the most frequent 1,000 AVL lemmas. The Reading/Writing and Vocabulary/Grammar subcorpora have the same pattern of AVL-lemma occurrence band frequencies as the whole corpus whereas in the Listening/Speaking subcorpus single occurrences are predominant for lemmas from any AVL band.

Further analysis indicates that even when AVL lemmas were shared between the Reading/Writing and Vocabulary/Grammar subcorpora (the two largest subcorpora), the average repetition rate was below 10 occurrences, the repetition rate necessary for recall vocabulary knowledge to develop incidentally. Nearly two thirds of the AVL lemmas shared among all subcorpora appeared 10 or more times in the materials.

Conclusion and future research

Findings are encouraging because they indicate that even when EAP lecturers do not make a conscious attempt to include high-frequency AVL lemmas when selecting and adapting their teaching materials, they do include many. However, the need for targeted vocabulary teaching became evident when we considered the repetition rate of AVL lemmas in the materials.

Since in-house materials are context specific, research is needed to examine academic vocabulary in materials used in EAP courses at other universities. In addition to examinations of written teaching materials, such studies should explore the role of academic vocabulary in teachers' materials selection and development and how teachers use these materials in the classroom.

In this study data were analysed quantitatively to identify AVL lemmas in the written materials and examine their repetition rate. Research into how many of

these AVL lemmas were the focus of direct teaching in the EAP materials and when this was the case and what kinds of vocabulary knowledge (e.g., meaning, collocations, grammatical properties) were targeted in vocabulary activities is necessary to provide a more thorough examination of how well EAP teaching materials cater towards students' academic vocabulary needs.

References

Alothman, K. 2014. Investigating the impact of focusing on academic vocabulary using multiple assessment measures. Unpublished PhD thesis, University of Roehampton.

Gardner, D., Davies, M. 2014. A new Academic Vocabulary List. Applied Linguistics 35, 305 –327.

Mazgutova, D., Kormos, J. 2015. Syntactic and lexical development in an intensive English for Academic Purposes programme. Journal of second language writing *29,* 3–15.

Uchihara, T., Webb, S., Yanagisawa, A. 2019. The effects of repetition on incidental vocabulary learning: a meta-analysis of correlational studies. Language Learning 69, 559–599.

The intonation of Italian verbless exclamatives

Patrizia Sorianello
Department of Letters, Language, Arts, University of Bari, Italy
https://doi.org/10.36505/ExLing-2020/11/0049/000464

Abstract

The present study aims to explore the prosodic properties of Italian verbless exclamatives (VEs), particular sentence structures without wh-quantifier and copula. A speech corpus formed by 250 VEs uttered by five subjects of a Southern variety of Italian was analyzed. The experimental results proved that VEs have a marked prosodic structure typically made up by two opposed constituents. The preposed predicative phrase is characterized by a salient intonation contour, while the grammatical subject is marginalized and shows a monotonous f0 pattern. The information structure is fixed too: the predicative constituent carries the new information, thus contrasting with the subject that expresses a given content.

Keywords: intonation, verbless exclamatives, Italian language

Introduction

Verbless exclamatives (from now VEs) are a particular type of sentence attested in diverse Romance languages used to express strong personal opinions. They are verbless configurations in which the ellipsis of the verb (normally the copula) produces reduced utterances with or without the wh-modifier, i.e. *Che buona questa torta! (What a good cake!)*, *Buona, questa torta! (Good, this cake!)*. This sentence structure shows a series of syntactic, information and prosodic restrictions. VE has a rigid syntactic order: the predicative complement precedes the grammatical subject of the clause since it is placed in the left periphery of the sentence. The order of the constituents is fixed and cannot be changed, **Questa torta, buona!* (**This cake, good!*). Moreover, in Italian the anteposed complement may be occupied only by gradable and evaluative adjectives, noun and adjective or, in alternative, by a noun preceded by an indefinite article i.e. *Una bontà, questa torta!* (*A goodness, this cake!*). If present, the initial wh-word is always 'che' (*what*), being impossible the presence of other wh-modifiers (**Come buona questa torta!*, **How good this cake!*), while the predicate must contain a personal emotional reaction, in line with the property of factivity, a peculiar semantic property of exclamatives. At the same time, the subject of the sentence holds demonstratives or deictic determiners, for this reason it does not allow indefinite lexical forms, **Buona, una torta!* (**Good, a cake!*).

As far as the information structure is concerned, VEs can never be considered neutral utterances. The predicative phrase conveys the information focus of the sentence, since it represents the unit which is informatively

relevant; in other words it expresses a widening and extreme-degree meaning, two intrinsic properties of exclamatives. On the contrary, the subject transmits a presupposed meaning.

There are diverse studies focused on the syntactic and semantic properties of VEs (Vinet, 1991; Gutierrez-Rexach 2001; Zanuttini & Portner 2003; Munaro, 2016), but the prosody of VEs has not received much attention so far. Narrowing the attention to the Italian language, some information can be deduced from Munaro (2016), while a detailed spectro-acoustic analysis is available in Sorianello (2017).

Methodology

This research intends to investigate the intonation contours of Italian VEs adopting an acoustical perspective. For this study we considered only VEs lacking of verb (silent copula) and of wh-word (silent *che*) in which the predicative complement has an adjectival nature i.e. *Interessante, questo libro!* (*Interesting, this book!*). The language under investigation is a Southern variety of Italian (Cosenza). Five participants aged between 22 and 30 (3 F, 2 M) were involved in a reading task. They produced 50 VEs embedded in short scenarios purposely constructed to simulate appropriate pragmatic contexts. Each target sentence was compared with both assertive (ASs) and wh- versions (wh-Es). The total corpus consisted of 250 VEs (50 sentences for 5 speakers), 250 wh-Es and 250 ASs (ie. *Buona, questa torta!\ Che buona è questa torta!\Questa torta è buona*). Audio recordings were acoustically analyzed with PRAAT by means of the following parameters: f0 maximum, f0 minimum, average f0, pitch range (in semitones), speech rate (sill/sec), duration of stressed vowels. Pitch accents and boundary tones were individuated on the f0 curve and labelled according to ToBI. Statistical significance was tested by Anova ($p < 0.05$).

Results and discussion

The intonation contours of all target sentences are falling (L%). Nevertheless, their realization is different. VEs have a final falling pitch pattern which does not overlap that of wh-Es or ASs, being characterized by a greater pitch range, a higher f0 initial value and a different intonational phrasing. The height of the initial f0 contributes to differentiate VEs from ASs; this phonetic cue plays a crucial role in the realization of a VE since it emphasizes the speaker-based evaluation. With reference to intonation, it is significant to observe that the pattern of VEs is formed by two opposed phases separated by an intermediate accent. The initial constituent is always salient: it shows a wide pitch range (in average 8 ST, SD: 2) and high values of intensity. The predicative complement is the only prominent portion of the sentence and receives the most salient pitch accent. In line with its expressive meaning, the preposed adjectival constituent presents an extra-high level (H* or L+H*) followed by a steep fall

that ends with a low intermediate accent (L-). On the contrary, the remaining part, syntactically coincident with the marginalized subject, always shows a low intonation contour (L* L%) and a compressed pitch excursion (in average 3 ST, SD: 1) without considerable movement within the stressed vowel; the intonation configuration is illustrated in Figure 1.

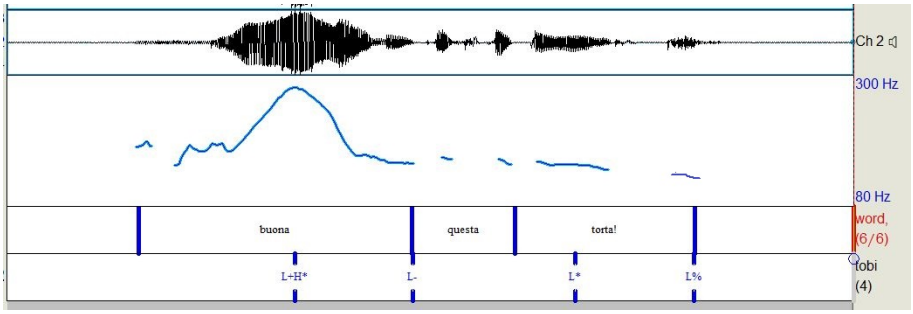


Figure 1. Waveform and f0 pattern of the VE *Buona questa torta!* realized by a female speaker.

As far as duration is concerned, the initial phrase is significantly longer in VEs than in control sentences; nuclear vowels are longer too, respectively VEs: 200 ms (SD: 38), wh-Es: 170 ms (SD: 35), ASs: 150 ms (SD: 22), all the statistical comparisons are significant ($p<.005$). This result is not surprising, in VEs the predicative complement, together with the deletion of the copula, causes a rhythmic and prosodic fracture. The adjective that conveys the focus of the exclamation shows a significant lengthening, a peculiar aspect that increases its unexpected content: the greater the sense of unexpectedness, the greater the degree of lengthening. The exclamatory force is not only given by the presence of longer nuclear vowels, but also by a general speech slowing; the results regarding speaking rate confirm this impression. VEs display a slow speaking rate, medially 4.8 sill/sec (SD: 0.5) whereas ASs and wh-Es reach the mean rate of 7 sill/sec (SD: 0.6) and 5.6 (SD: 0.7) respectively; the differences are statistically significant ($p <.005$).

The information structure of VEs is rigid too. It can be represented as follows:

Buona $_{NEW}$ *questa torta!* $_{GIVEN}$

The order is fixed: the predicative complement coincides with the NEW information, the subject with the GIVEN one. Prominence is the result of the interaction of several factors. The information focus receives prosodic salience; all the acoustic parameters are activated in an 'extra' way making the predicative constituent longer, more intense and with a greater pitch excursion. All this ensures the realization of the widening meaning of the exclamative contributing, at the same time, to the achievement of an adequate exclamatory

force. The adjectival constituent conveys the surprising and emotional content of the sentence, since it expresses an evaluation at an extreme point of a scale. On the contrary, the subject is marginalized and represents old information; it refers to something which is already present in the common ground of the participants. Accordingly, it is produced with no accentuation: a flat intonational contour (L* or !H+L*) placed on the lowest f0 range characterized it. The boundary between the two constituents is obligatory signalled by a low phrase accent (L-). In general, information and prosodic structures establish a strong relationship: the predicative phrase and the subject one show in fact a polarized behaviour for duration, speaking rate, intensity, pitch range and intonation contour. The marked position of the predicative constituent has deep effects on the other levels. It transmits a high degree of emotional arousal, but also a new information. The surprising content is limited only to the first part of the VE, a condition reinforced also by the lack of both the wh-word and the copula; as a consequence, the final region of the utterance is totally deaccented.

This study added new data to the prosodic description of VEs. The analysis showed that in Italian VEs are rigid sentential structures formed by two constituents distinguished by opposite information and prosodic features. This peculiar organization is perfectly congruent with their pragmatic and expressive content.

References

Gutiérrez-Rexach, J. 2001. Spanish exclamatives and the interpretations of the left periphery. In D'Hulst, Y., Rooryck, J. & Schroten, J. (eds.), Romance languages and linguistic theory 1999, Amsterdam/Philadelphia, J. Benjamins, 167-194.

Munaro, N. 2016. Verbless predicative structures across Romance. *Journal of Linguistics* 52, 609-637.

Rett, J. 2008. A degree account of exclamatives. In Friedman T. & S. Ito (eds.) *Proceedings of SALT XVIII*, Ithaca, New York, CLC Publications, 601-608.

Sorianello, P. 2017. Sulla realizzazione prosodica delle esclamative nominali. Rivista Italiana di Dialettologia 41, 147-170.

Vinet, M. T. 1991. French non-verbal exclamative constructions. Probus 3, 77-100.

Zanuttini, R., Portner, P. 2003. Exclamative clauses: at the syntax-semantics interface. Language 79, 39-81.

Automated speech analysis enables MCI diagnosis

Charalambos Themistocleous[1], Marie Eckerström[2], Dimitrios Kokkinakis[3]

[1]Department of Neurology, Johns Hopkins University, USA
[2]Department of Psychiatry & Neurochemistry, University of Gothenburg, Sweden
[3]Department of Swedish, University of Gothenburg, Sweden

https://doi.org/10.36505/ExLing-2020/11/0050/000465

Abstract

Mild Cognitive Impairment (MCI) is a condition characterized by cognitive decline greater than expected for an individual's age and education level. In this study, we are investigating whether acoustic properties of speech production can improve the classification of individuals with MCI from healthy controls augmenting the Mini Mental State Examination, a traditional screening tool, with automatically extracted acoustic information. We found that just one acoustic feature, can improve the AUC score (measuring a trade-off between sensitivity and specificity) from 0.77 to 0.89 in a boosting classification task. These preliminary results suggest that computerized language analysis can improve the accuracy of traditional screening tools.

Keywords: acoustic analysis, machine learning, cognitive impairment, MMSE.

Introduction

Mild Cognitive Impairment (MCI) is a syndrome characterized by cognitive decline greater than expected for an individual's age and education level. The Mini-Mental State Exam (MMSE) is a screening tool for cognitive impairment often employed for evaluating individuals for MCI. Although, it has been considerably validated and is extensively used, it has received criticism as not being effective in detecting MCI. In an earlier work, Fraser, Lundholm Fors, Eckerström, Themistocleous, and Kokkinakis (2018), examined the utility of augmenting MMSE scores with automatically extracted linguistic information from a narrative speech task to better differentiate between individuals with MCI and healthy controls in Swedish. The study found that with the addition of just four linguistic features, the AUC score (measuring a trade-off between sensitivity and specificity) is improved from 0.68 to 0.87 in logistic regression classification. These results suggested that the accuracy of traditional screening tools may be improved through the addition of computerized language analysis. In this study, our goal is to determine whether cognitive decline can be estimated using acoustic information, as we have shown that speech production

is impaired in patients with MCI (Themistocleous, Eckerström, & Kokkinakis, 2018, 2020; Themistocleous, Kokkinakis, Eckerström, Fraser, & Fors, 2018). Thus, contributing to the work of identifying automatic diagnostic markers that have the potential to facilitate clinical evaluation and therapy.

Methodology

Participants for this study were recruited from the Gothenburg MCI study (Wallin, Nordlund, Jonsson, & others, 2016) a large-scale longitudinal study that aims to advance the nosological understanding in AD and other types of dementia (Wallin et al., 2016). Additional assessment tests were conducted for the purposes of the Riksbankens Jubileumsfond – The Swedish Foundation for Humanities & Social Sciences "Linguistic andextra-linguistic parameters for early detection of cognitive impairment" research grant (NHS 14-1761:1) where speech recordings from the cookie theft picture description task were elicited. Thirty healthy controls and 25 MCI—between 55 and 79 years old (M=69, SD=6.4) participated in the study (see Table 1). The two groups did not differ with respect to age [t(52.72)= −1.8178,p=n.s.] and gender (W=1567.5, p=n.s.) significantly. The recordings were conducted in an isolated environment at the University of Gothenburg. The recordings were analyzed acoustically, using advanced acoustic analysis and signal processing algorithms (see for the methodological analysis (Themistocleous et al., 2020)).

For the speech part, participants were asked to describe what they could see in the "Cookie Theft" picture from the Boston Diagnostic Aphasia Examination (Goodglass, Kaplan, & Barresi, 2001), widely employed picture that aims to elicit narrative speech. The picture description task was audio recorded using a Zoom H4N audio recorder, and the audio recordings were analyzed acoustically using the open source software for acoustic analysis Praat (Boersma & Weenink, 2018). Specifically, we analyzed speech sounds and measured acoustic properties related to voice quality and speech fluency. Measurements of the number of syllables, number of pauses, average syllable duration, phonation time, speech rate, articulation rate, and speaking time. We employed a machine learning approach to classifying patients with MCI from HC. We trained two Boosting Classification models (i) a Base Boosting Classification model with the MMSE, speaker age, gender, and education and (ii) an Augmented Boosting Classification Model that includes the added acoustic measures. Both Boosting Models add predictors to a decision tree ensemble, each one correcting its predecessor. Boosting fits the new predictor to the residual errors made by the previous prediction, instead of modifying the weights for every incorrectly classified observation.

Results

The estimated accuracy of each run is shown in Figure 1. The Base Boosting Model provided 75% classification accuracy and the Augmented Boosting

Model had an increased accuracy, namely 79%. These measures are reflected in the precision and recall of the two models (shown in Table 1). Figure 1 shows the AUC of the Based Boosting Model and the Augmented Boosting Model, the AUC for the MCI was increased in the Augmented Boosting Model from 76% to 89%. The Based Boosting Model employed primarily MMSE and education whereas the Augmented Boosting Model employed Articulation Rate in addition to MMSE, Education, and Gender to perform the classification task. The relative influence of the predictors of the Base Bosting Model and the Augmented Boosting Model are shown in Figure 1. In addition to the MMSE and education, the articulation rate and Gender affect the Augmented Boosting Model.

Table 1. Evaluation Metrics of the Base Boosting Classification model and the Boosting / Acoustics Classification model.

		Precision	Recall	F1 Score	AUC
Base Boosting	HC	0.647	0.917	0.759	0.802
	MCI	**0.909**	**0.625**	**0.741**	**0.766**
	Average / Total	0.797	0.750	0.748	0.784
Aug. Boosting	HC	0.800	0.800	0.800	0.995
	MCI	**0.769**	**0.769**	**0.769**	**0.887**
	Average / Total	0.786	0.786	0.786	0.941

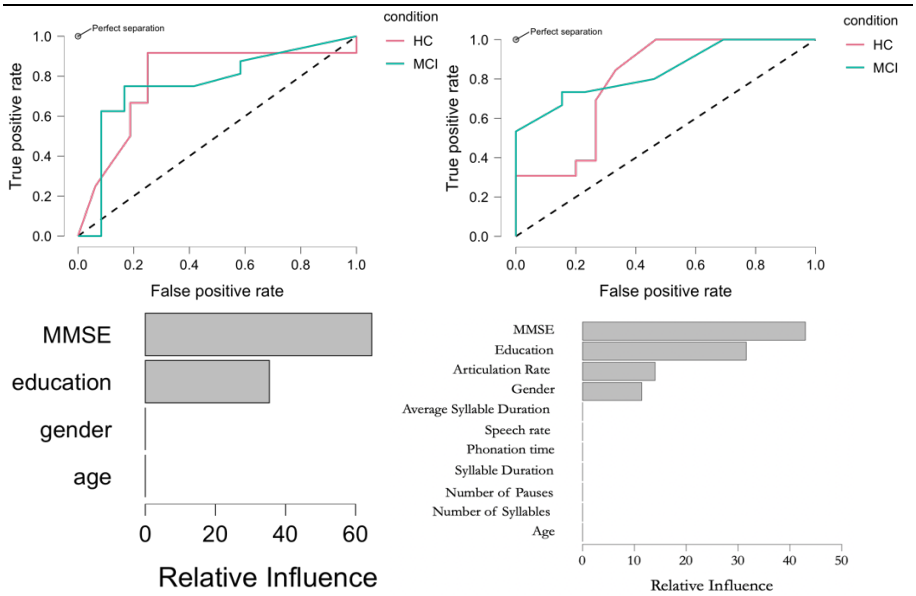


Figure 1. Shows the ROC curve plot for the Base Bosting Model (Panel A) and Boosting Model / Acoustics (Panel B) and relative influence of predictors of the Based Boosting Model (Panel C) and the Augmented Boosting Model (Panel D).

Discussion

In this study, we aimed to determine the use of acoustic measures to improve MCI classification, relative to using MMSE, education, and age scores alone. The results were positive, showing that the Augmented Boosting Model had an improved AUC over the Base Boosting Model by including just one acoustic measure. In the Augmented Boosting Model, the MCI AUC was improved from 0.76 in the Base Model using MMSE, age, gender and education to 0.89 by allowing the classifier to include the Articulation Rate in addition to the MMSE score. Speech contains information about the cognitive state of an individual and can reveal impairment in individuals with MCI as was shown in several earlier studies. This study provides promising results for automated measures of cognitive decline. Automated diagnostic measurements can be conducted at primary care centers and memory clinics quickly and easily assisting the work of clinicians, speech and language therapists and caregivers. The main limitation of the study is that we considered mainly temporal acoustic measures, as these are known to be affected in MCI, yet in our future research we aim to include measures of voice quality and prosody as well. Moreover, we plan to combine these measures with measures of grammar and discourse, as the Cookie Theft picture description task reduces the possibilities of eliciting a variability of linguistic expressions manifested using phonetic and grammatical means, such as various types of speech acts, questions, etc. We expect that tasks eliciting discourse and conversation will enable the assessment of linguistic and cognitive skills of individuals. Nonetheless, this study provides promising results towards augmenting and automating traditional evaluation tests using computational techniques from computational linguistics and machine learning.

References

Boersma, P., Weenink, D. 2018. Praat. Retrieved from http://www.praat.org

Fraser, K., Lundholm Fors, K., Eckerström, M., Themistocleous, C., Kokkinakis, D. 2018. Improving the sensitivity and specificity of MCI screening with linguistic information. LREC, Miyazaki, Japan.

Goodglass, H., Kaplan, E., Barresi, B. 2001. BDAE-3: Boston Diagnostic Aphasia Examination–Third Edition: Lippincott Williams & Wilkins Philadelphia, PA.

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. doi:10.3389/fneur.2018.00975

Themistocleous, C., Eckerström, M., Kokkinakis, D. 2020. Voice quality and speech fluency distinguish individuals with Mild Cognitive Impairment from Healthy Controls. PLoS One, 15(7), e0236009. doi:10.1371/journal.pone.0236009

Themistocleous, C., Kokkinakis, D., Eckerström, M., Fraser, K., Fors, K. L. 2018. Effects of Cognitive Impairment on vowel duration. ExLing 2018, 113.

Wallin, A. et al. 2016. The Gothenburg MCI study. Journal of Cerebral Blood Flow & Metabolism, 36(1), 114-131. Retrieved from https://doi.org/10.1038/jcbfm.2015.147.

Morphological awareness in L1 and L2 reading skills

Vassiliki Tsela[1], Georgia Andreou[1], Maria Liakou[2], Julie Baseki[2]
[1]Department of Special Education, University of Thessaly, Greece
[2]Hellenic Open University, Greece
https://doi.org/10.36505/ExLing-2020/11/0051/000466

Abstract

The present study investigated the effect of morphological awareness on three measures of reading, namely decoding, fluency and comprehension, in Greek as a first language (L1) and in French as a foreign language (L2). Morphological awareness was assessed via two tasks, a verb inflection task and a word production task. The results of this study indicated that the student's performance in the two morphological tasks was significantly associated with their performance in the reading tasks. Our results support our hypothesis that morphological awareness can be a significant predictor of the high or low performances in decoding, reading fluency and reading comprehension in both L1 and L2 and it plays a critical role in reading efficiency.

Keywords: morphological awareness, reading skills, Specific Reading Disorder, Greek (L1), French (L2)

Introduction

The relationship between morphological awareness and reading skills has been repeatedly documented. The contribution of morphological awareness to both decoding and comprehension reading skills increases with age and consequently, it gets more important in later reading development stages (Carlisle, 2000; Tong, Deacon, Kirby, Cain, & Parrila, 2011). Furthermore, research has shown that students with Specific Reading Disorder (SRD) have limited morphosyntactic skills when compared to students with Good Reading Skills (GRS) in both their first (L1) and second/foreign (L2) language (Casalis, Colé, & Sopo, 2004).

Purpose of the study

The aim of the present study was to investigate the effect of morphological awareness on three measures of reading, namely decoding, fluency and comprehension, in Greek as a first language (L1) and in French as a foreign language (L2). For both students with GRS and students with SRD, morphology skills are expected to predict students' reading skills in each language.

Method

The sample consisted of two groups of students in the last two grades of primary school (Grades 5 and 6), students with GRS (N=49) and students with SRD (N=49). Morphological awareness was assessed via two types of tasks, a verb inflection task and a word production task, from "Learning Disabilities Reading Inventory" (Padeliadu & Antoniou, 2008), a standardized reading assessment tool for the Greek language, and a reading assessment tool for the French language developed for the purposes of the present study. Pearson's r correlation coefficient was used for the statistical analysis of the data.

Results

The results of the study revealed significant correlations between morphological awareness and reading skills in both languages in both groups.

Table 1. Correlation between morphological awareness and reading skills in Greek and in French in GRS and SRD students.

Task	Group/Language	Decoding	Fluency	Comprehension
Verb inflection/Greek	GRS/Greek	*r*=0,684 *P*<0,001*	*r*=0,498 *P*<0,001*	*r*=0,318 *P*<0,001*
	SRD/Greek	*r*=0,306 *P*<0,001*	*r*=0,335 *P*<0,001*	*r*=0,475 *P*<0,001*
Verb inflection/French	GRS/French	*r*=0,467 *P*<0,001*	*r*=0,457 *P*<0,001*	*r*=0,537 *P*<0,001*
	SRD/French	*r*=0,719 *P*<0,001*	*r*=0,517 *P*<0,001*	*r*=0,693 *P*<0,001*
Production of compound words/Greek	GRS/Greek	r=0,746 *P*<0,001*	*r=0,*563 *P*<0,001*	*r*=0,416 *P*<0,001*
	SRD/Greek	*r*=0,258 *P*=0,007*	*r*=0,353 *P*<0,001*	*r*=0,449 *P*<0,001*
Word production/French	GRS/French	*r*=0,550 *P*<0,001*	*r*=0,513 *P*<0,001*	*r*=0,625 *P*<0,001*
	SRD/French	*r*=0,482 *P*<0,001*	*r*=0,611 *P*<0,001*	*r*=0,665 *P*<0,001*

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

Discussion

An important finding of the present study is the high correlation between morphological inflection and decoding as well as fluency. It seems that the functional character of inflection suffixes in the verb inflection task influences the precision and the fluency in reading both in Greek and in French for both

groups of students, who have completed the first stage of learning to read and write in their mother tongue. More specifically, inflection suffixes bear phonological information (Mahony, Singson, & Mann, 2000; Ralli, 2005) which can have an impact on word decoding, as the phonological information found in some verb suffixes influences the word pronunciation and causes changes in the intonation and the pronunciation of the radical part of the word both in Greek and in French. Moreover, the high correlation between morphological awareness and word decoding and fluency may be linked to the complex morphological structure of some of the words which were used in the tasks. The high correlations between performance in the tasks of morphological inflection and reading comprehension show that text comprehension is influenced by morphological inflection awareness, since inflectional suffixes bear morphosyntactic and semantic information.

As for the students' morphological awareness performance on the word production tasks, the present study showed that there is strong correlation between performance on these tasks and decoding and fluency in both languages. These findings indicate that the phonological, syntactic and semantic information included within production suffixes influence word decoding (Mahony, Singson, & Mann, 2000; Ralli, 2005). The morphological production tasks included items in which the derived word was phonologically different from the radical. This condition of phonological change, which caused difficulties to the students with SRD, shows the significant dependence of phonological skills on morphological processing and, consequently, the impact of morphological awareness on reading decoding. Additionally, the derivatives and the compound suffixes in the Greek language are semantically opaque forms (Ralli, 2005) and can cause difficulties to students with SRD. The same happens with derived words in the French language. In conclusion, low performance in reading, decoding and fluency, as a consequence of phonological deficit, as well as the phonological and orthographic changes between radical and derived words seem to influence and thus explain the low performances of students with SRD in morphological awareness tasks.

The results of the present study also showed that there is strong correlation between the students' performance on the morphological production task and the reading comprehension task in both languages for both groups. It is quite interesting that morphological production is a significant predictive indicator of reading skills for upper grades of primary school, when reading has started to be an automatized skill. According to our hypothesis, the more the decoding ability improves, with the number of morphologically complex words increasing in upper grades, the more morphological awareness through the recognition of words bearing familiar morphemes and, consequently, fluent reading is facilitated. Furthermore, phonological and semantic transparency between derivatives and radicals, as well as between compound words and their components, seem to facilitate the interpretation of unknown words and,

therefore, influence reading comprehension (Kuo & Anderson, 2006). Regarding Greek students in the early stages of learning French as a foreign language, the present study showed that morphological awareness skills can predict reading comprehension skills.

For the students with SRD, poor performance in morphological awareness should be expected, since morphemes also bear phonological characteristics and, consequently, the difficulty in phonological processing which leads to the difficulty in decoding and fluency seems to be positively related with morphological skills, as well. It seems that deficient skills in phonological processing are responsible for the limited skills in morphological processing for students with SRD, since the morphological awareness tasks contained phonologically opaque words.

It is concluded that morphological skills can predict reading performance for both students with SRD and students with GRS in both their L1 and L2. Moreover, the correlations between morphological skills and reading skills found in this study indicate that morphological skills can play a role in the reading performance in an alphabetic language with a rich morphology like the Greek language, but also in the reading performance in a language with a simpler morphology, like French.

References

Carlisle, J.F. 2000. Awareness of the structure and meaning of morphologically complex words: Impact on reading. Reading and writing, 12(3), 169-190.

Casalis, S., Colé, P., Sopo, D. 2004. Morphological awareness in developmental dyslexia. Annals of Dyslexia, 54(1), 114-138.

Kuo, L., Anderson, R.C. 2006. Morphological awareness and learning to read: a cross-language perspective. Educational Psychologist, 41(3), 161-180.

Mahony, D., Singson, M., Mann, V. 2000. Reading Ability and Sensitivity to Morphological Relations. Reading and Writing: An Interdisciplinary Journal, 12, 191-218.

Padeliadu, S., Antoniou, F. 2008. Learning Disabilities Reading Inventory. Athens: YPEPTH: EPEAEK.

Ralli, A. 2005. Morphology. Athens: Pataki Publishers.

Tong, X., Deacon, S.H., Kirby, J.R., Cain, K., Parrila, R. 2011. Morphological awareness: A key to understanding poor reading comprehension in English. *Journal of Educational Psychology, 103*(3), 523–534.

Vowel length of Czech trisyllabic words in L1 Russian speakers

Jitka Veroňková[1], Giorgi Gersamia[2]

[1]Institute of Phonetics, Charles University, Czech Republic

[2]Institute of Czech Language and Theory of Communication, Charles University, Czech Republic

https://doi.org/10.36505/ExLing-2020/11/0052/000467

Abstract

Languages differ in how vowel length is applied, which can affect the acquisition of a second language. In Czech, the length is phonological. There are practically no restrictions on its occurrence. It is also completely independent of word stress. In Russian, the length of vocals does not have a phonological status but can be an accompanying characteristic in the implementation of a word stress. The subject of the experiment is the perceptual analysis of vocal quantity in Czech as L2 in Russian speakers (8 subjects). The material consists of recordings of a set of trisyllabic words (48 lexemes, 256 items), in which the structure of quantity (8 different patterns) is checked.

Keywords: Czech as L2, Russian as L1, vowel length, perception, trisyllabic words.

Introduction

The vowel length belongs among the segmental phenomena causing difficulty in production and perception for speakers of Czech as L2, including Russian L1 speakers. (Romaševská & Veroňková 2016) In Russian, vowel quantity is present; however, it does not have phonological status, and it is associated with word stress. On the contrary, in Czech, the vowel length is phonological, and it is entirely independent of Czech word stress; the stressed syllable can be both long and short. A long syllable is not tied to a specific position in a word, and words may contain more than one long vowel; on the other hand, a word may contain no long vowel. Incorrect pronunciation of vowel length may result in the change of meaning of a message, or at least it may contribute to worse intelligibility of speech (in the sense of Munro & Derwing 1995).

In this paper, the L1 Czech listeners' perception of the vowel length of L2 Czech speakers with Russian mother tongue in trisyllabic words with various length patterns is presented.

Methodology

Short (S) and long (L) vowels in a trisyllabic word may be combined in eight different ways (SSS, LSS, SLS... LLL), with all these patterns attested in Czech. For each pattern, we selected six words (2 nouns, 2 adjectives, 2 verbs), i.e.,

48 target words altogether. When collecting the suitable words, an attempt was made to apply the parameter of similarity with Russian, but it was not possible to establish definite criteria for determining which Czech word has a similar equivalent in Russian and which does not.

A story was created using the target words, which were placed in the middle of the sentence. In order to reduce the influence of the surroundings on the realization of the vowel quantity, the surrounding words contained only short vowels and were at least disyllabic in order to avoid including the target word with a surrounding word in one stress group.

11 female speakers with Russian as L1 – students of Czech language courses who were preparing to study at Czech universities (level B1–B2 according to CEFR) – were recorded. Recordings were taken individually in a sound-treated room (AKG C 4500 B-BC microphone, sample rate 32 kHz, 16-bit depth), and speakers were provided a short time for preparation.

We performed a perceptual analysis using Praat software (Boersma & Weening 2019) and determined vowel lengths in the target words and the position of word stress. In case of uncertainty, the listening was performed by another native listener, phonetically educated.

Results

Of the 528 items analysed, 16.3% were excluded because of slips of tongue or repetitions. A relatively larger number of excluded words belonged to three of four patterns containing at least two long vowels (see Figure 1, left column). One-quarter of LLS words (25.8%) and one- fifth of LSL and LLL words (19.7% both) were excluded. We believe that the occurrence of more than one long vowel in a word may have contributed to the pronunciation difficulties.

The successful pronunciation of vowel length was achieved in 37.3% of items, i.e., in less than half of them. However, the success score varies considerably, regarding particular patterns (see Figure 1, right column). The highest success score was achieved in the LSS words (77.2%), with a long vowel in the first syllable (which regularly bears word stress in Czech). None of the other patterns achieved a success rate of 50%. The lowest scores were noticed in the two patterns mentioned above with the highest number of excluded items – LLS 10.2% and LLL 17.0%, containing at least two long vowels. The SSS pattern also displayed a very low success score (25.9%), which, unlike these patterns, contained only short vowels. When pronouncing SSS words, speakers prefered to use SLS pattern, often with word stress on the second syllable, i.e., long in the production.

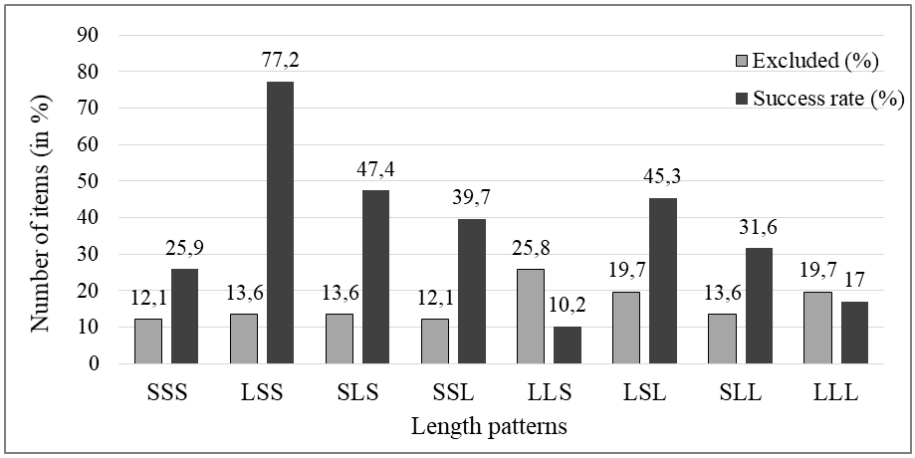


Figure 1. Number of items excluded from the analysis and successfully pronounced, categorized by length pattern. S – short vowel, L – long vowel.

Regardless of whether it was the correct realization or not, speakers preferred LSS and SLS patterns in production. Almost 60% of all the realizations are covered by these two patterns. On the contrary, the LLS and LLL patterns were rarely used in the production (each only in 2% of realizations). See Figure 2.

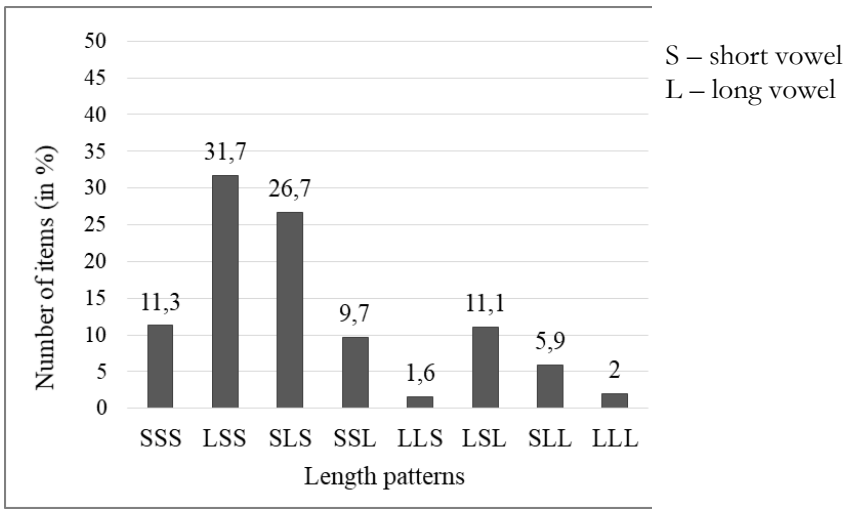


Figure 2. Number of pronounced variants categorized by length patterns.

Conclusion

It has been confirmed that the Russian speakers truly have difficulty pronouncing short/long vowels in Czech. The type of vowel length pattern, especially the accumulation of lengths, may affect the success of pronunciation. Two successive long vowels proved the most difficult, with their combination in the word-final position (SLL) being less difficult than in the word-initial position (LLS), or in the case of three long vowels (LLL). It is apparent that the position of long vowels may influence pronunciation success as well. The pattern LSL with two long vowels separated by a short vowel, representing a relatively more successful pattern, is another example of that. Nevertheless, the difficulties are not only caused by long vowels, but also by words containing canonically only short vowels; in such cases the production of the long vowel may be motivated by the word stress. The relationship between long vowels and word stress also explains the preference of the LSS pattern with a long vowel in the first, stressed syllable in Russian speakers' production. Similarly, the wider use of SLS, in which speakers implemented the word stress mainly on the second syllable, can be explained. It is evident that Czech vowel quantity practice in Russian speakers should be combined with word stress exercises. Familiarity with patterns causing more difficulties has an impact on the creation of pronunciation exercises that may be focused on the specific combinations.

Acknowledgements

This research was supported by the Czech Science Foundation project No. 18-18300S "Phonetic properties of Czech in non-native and native speakers' communication". We would like to thank Petra Poukarová (Faculty of Arts, Charles University, Prague) for contributing to the listening analysis.

References

Boersma, P., Weenink, D. 2016. Praat: Doing Phonetics by Computer (Version 6.0.25). http://www.praat.org.

Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR), https://www.coe.int/en/web/common-european-framework-reference-languages.

Munro, M.J., Derwing, T.M. 1995. Foreign Accent, Comprehensibility, and Intelligibility in the Speech of Second Language Learners. Language Learning 49(1), 73–97.

Romaševská, K., Veroňková, J. 2016. How Czech speech of Russian-speaking learners is perceived by native speakers of Czech and its correlation with age factors and language competence. In Besters-Dilger, J., Gladkova, H. (eds.), Second Language Acquisition in Complex Linguistic Environments. Russian Native Speakers Acquiring Standard and Non-Standard Varieties of German and Czech, 147–176, Frankfurt am Main, Peter Lang.

On the acquisition of Greek free choice items

Evangelia Vlachou[1], Dimitrios Kotopoulis[2], Spyridoula Varlokosta[2]
[1]Department of French, National and Kapodistrian University of Athens, Greece
[2]Department of Linguistics, National and Kapodistrian University of Athens, Greece
https://doi.org/10.36505/ExLing-2020/11/0053/000468

Abstract

Children acquire quite late scalar implicatures associated with quantification and have the tendency to interpret existential quantifiers as universals (e.g., Smith, 1980; Noveck, 2001; Papafragou and Musolino, 2003). Free choice Items (FCIS) are also associated with scalar implicatures depending on whether they are full set or subset FCIs (e.g., Vlachou 2012, 2020). This paper presents experimental results showing that 9-, 10- and 11-year-old children and adults perform better on full set than on subset FCIs. It is argued that adults perform better than children in sentences with subset FCIs as the "not-all" pragmatic inference is acquired late. Difficulties in sentences with subset FCIs in adults are due to absence of domain alternatives.

Keywords: free choice items, scalar implicatures, quantification, alternatives, semantics

Introduction

The literature on the acquisition of quantifiers (e.g., Noveck, 2001; Papafragou and Musolino, 2003; Smith, 1980) has put forward the hypothesis that children acquire quite late scalar implicatures associated with quantification and have the tendency to interpret existential quantifiers as universals: the "not-all" meaning component of existential quantifiers is acquired only at a later stage.

The close relation between quantification and Free Choice Indefinites (FCIs) such as *any* has been extensively documented (see Dayal, 1998 and Kadmon and Landman, 1993 inter alia).

(1) a. Any student can solve this problem.
    b. All students can solve this problem.

(2) a. I did not talk to just any student.
    b. I did not talk with some student or other.

Vlachou (2012, 2020) argued that, the core semantic property that defines FCIs as a class and distinguishes them from other lexical items is that they have descriptive contents that, except for referring to alternatives, bring information on these alternatives. They form two different interpretational categories: full set, associated with universal quantification over alternatives (an "all" meaning) and subset FCIs, associated with existential quantification over alternatives (a "not-all" meaning). Full set FCIs (*opjosdhipote* (*o-dhipote*, hereafter) require that all

alternatives, without exception, are considered. Subset FCIs (subset 1: *enas/o opjosdhipote* (*enas/o o-dhipote*, hereafter); subset 2: *opjos opjos* (*o- o-*, hereafter)) restrict the domain of quantification of the alternatives to one of its subsets.

The difference in the acquisition of universal and existential quantifiers makes one predict that the acquisition of full set and subset FCIs differs, as well. Moreover, recent accounts, as in Tieu et al. (2015), have shown that the acquisition of scalar implicatures becomes harder when relevant alternatives are not provided by the context. To test both predictions, we set up an experiment to investigate the tendencies in the acquisition of Greek FCIs.

Methodology

Participants

The sample consisted of 154 participants: 120 children – 40 9-year-old attending 4th grade (age range: 9;6 - 10;5; mean age: 9;11), 40 10-year-old attending 5th grade (age range: 10;6 - 11;5; mean age: 10;11), and 40 11-year-old attending 6th grade(age range: 11;6 - 12;5; mean age: 11;11) – and a control group of 34 adults. All participants were monolingual native speakers of Greek.

Materials

We developed two written tasks, assessing production (*Task A*) and comprehension (*Task B*) of FCIs. *Task A* was a sentence completion task that contained 60 sentences with gaps: 30 sentences had a full set FCI as target item (3) and 30 sentences a subset FCI (4). *Task B* contained 40 pairs of sentences. In each pair, the two sentences were identical, except for the FCI they contained; one always had a full set FCI (5a), while the other a subset FCI (5b).

(3) Tha eprepe na kseris oti egho s'aghapo perisotero apo opjondhipote alo.
'You should know that I love you more than anybody else.'

(4) Egho dhen ime (enas/o opjosdhipote / opjos ki opjos). Ime o kaliteros su filos.
'I'm not just anyone. I am your best friend.'

(5) a. Opjosdipote mathitis echi tsanda. (target)
'Any pupil has a schoolbag.'
b. #?Enas opjosdhipote mathitis echi tsanda.

Sentences in both tasks were fairly short and deprived of (long) pragmatic context in order to examine how participants produce and comprehend "out of the blue" sentences with FCIs. If not, our experiment would corroborate the validity of the hypothesis that pragmatic context is relevant to the understanding of FCIs.

Results

Significant differences were observed among the four age groups in both tasks regarding the mean scores of successful performance (see Table 1). As expected, fourth graders scored lower than the other three groups, while adults exhibited the highest mean scores. In Task A, a significant difference arose between the performance of fourth graders and that of all other groups, while in Task B, the adults' performance differed significantly from that of all the other groups (Table 2).

Table 1. Mean scores and statistics for the four age groups in the two tasks.

Group	Task A Mean (Range)	Task B Mean (Range)
4th grade	34.93 (8-54)	24.68 (16-33)
5th grade	41.25 (21-55)	25.7 (14-39)
6th grade	41.8 (22-56)	27.43 (18-35)
adults	44.41 (30-52)	31.65 (24-38)
Kruskal-Wallis	$\chi2 = 23.51$, df = 3, p = .000	$\chi2 = 49.69$, df = 3, p = .000

Table 2. Statistics for pair-wise between-groups comparisons.

	TASK A		TASK B	
Comparisons	Test Value (U)	p-value	Test Value (U)	p-value
4th vs. 5th grade	467	.001	703.5	.35
4th vs. 6th grade	461	.001	582	.03
4th vs. adults	264	.000	98.5	.000
5th vs. 6th grade	759	.69	681.5	.25
5th vs. adults	498	.04	156.5	.000
6th vs. adults	562.5	.2	236	.000

In both tasks, the overall difference in correct answers between full set and subset FCIs yielded a significant result (Task A: Z = -9.051, p = .000; Task B: Z = -10.448, p = .000). The subsequent testing of differences between the two types of FCIs within each group for each task showed that all four groups performed better in full set than in subset FCIs (Table 3).

Table 3. Means and statistics for full set vs. subset comparisons per group.

	TASK A			TASK B		
	Means		Statistics (Z; p)	Means		Statistics (Z; p)
Group	Full Set	Subset		Full Set	Subset	
4th grade	25.1	9.83	-5.159; .000	14.97	7.68	-5.129; .000
5th grade	26.3	14.95	-5.352; .000	15.85	2.06	-5.521; .000
6th grade	25.22	16.57	-4.189; .000	15.15	9.65	-5.283; .000
Adults	24.94	19.47	-2.608; .009	17.50	11.79	-5.046; .000

Discussion and conclusions

Our prediction that full set FCIs are acquired earlier than subset FCIs was born out by the results of this study. Our experiment revealed that 9-, 10- and 11-year-old children have a clear tendency of using and comprehending sentences with full set FCIs over sentences with subset FCIs. Moreover, the acquisition of sentences with subset FCIs becomes better as age progresses but is not adequate yet at the age of 11. Subset and full set FCIs follow different acquisitional paths, as they have different semantic features. Subset FCIs are associated with a "not-all" implicature, known to cause problems to children. Full set FCIs, on the other hand, are the default option for 9-, 10- and 11-year-old children, as they are associated with universal quantification over alternatives, known not to cause any problem in the acquisition of quantification.

Our experiment leads us to another interesting finding that puts our study in line with studies that have detected difficulties in the comprehension and production of sentences with "not-all" implicatures in adults too (e.g., Tieu et al. 2015). Even adults do not perform well on scalar implicatures when they are unable to retrieve the alternatives in question. Once alternatives are provided, they both comprehend and produce sentences with scalar items. Returning to our experiment, data contained fairly short sentences deprived of any pragmatic context or any picture that would enable participants to retrieve the alternatives in question. Thus, our results serve as further evidence for the validity of the hypothesis that once alternatives are not provided by the context, even adult participants have difficulties computing sentences with FCIs.

References

Dayal, V. 1998. Any as inherently modal. Linguistics and philosophy 21, 433-476.

Kadmon, N., Landman, K. 1993. Any. Linguistics and philosophy 16, 353-422.

Noveck, I. 2001. When children are more logical than adults: experimental investigations of scalar implicature. Cognition 78, 165-188.

Papafragou, A., Musolino, J. 2003. Scalar implicatures: experiments at the semantics-pragmatics interface. Cognition, 86(3), 253-282.

Smith, C. 1980. Quantifiers and question answering in young children. Linguistic investigations. Journal of experimental child psychology 40, 191-205.

Tieu, L., Romoli, J., Zhou, P., Crain, St.. 2016. Children's knowledge of free choice inferences and scalar implicatures. Journal of semantics 33(2), 269-298.

Vlachou, E. 2012. Delimiting the class of free choice items in a comparative perspective: evidence from the database of the French and Greek free choice items. Lingua 122(14), 1523-1568.

Vlachou, E. 2020. Indéfinis et implicatures: le cas de n' importe *qu-* et *un N quelconque*. In C. Beyssade (ed.) Implicature. ISTE.

L2 vocabulary learning motivation by Chinese EFL learners

Xuan Wang
Department of Languages, Literatures and Linguistics, Syracuse University, US
https://doi.org/10.36505/ExLing-2020/11/0054/000469

Abstract

L2 motivation has been proved by a plethora of studies to positively affect various domains of L2 learning, among which L2 vocabulary learning is relatively underexplored in the literature. This study, therefore, explores the characteristics of L2 vocabulary learning motivation by Chinese EFL learners and investigates how motivation relates to self-regulated learning strategies. The study employs a mixed-method approach with 47 Chinese EFL learners. Two instruments, the motivation questionnaire and the learning strategies questionnaire were employed, and ten participants were interviewed with regard to self-regulated vocabulary learning. The results reveal the instrumentality (i.e. promotion and prevention) of Chinese EFL learners in vocabulary learning, which is related to their learning strategies and self-regulation.

Keywords: L2 Motivation; Vocabulary Learning, Self-regulation; Learning Strategies

Theoretical background

L2 motivation, as an intricate construct of individual differences, plays a vital role in the L2 learning performance and learning outcome (Dörnyei & Ushioda 2009). Over the last three decades, motivation studies underwent three main phrases of research development (Boo et al. 2015). The first phase of motivation studies centred on the social-psychological perspective of language learning; during this phrase, Gardner (1985) proposed the seminal dichotomy of integrative and instrumental motivation. The second phase focuses on the self-determination theory (Deci & Ryan 1985) from the perspectives of cognitive and educational psychology, according to which, intrinsic (e.g., internal enjoyment and interest) and extrinsic motivations (e.g., external rewards) are two constructs of L2 motivation. The third phase is a key period in which most current motivation studies were conducted; it emphasizes the contextual and dynamic features of learner motivation in relation to the L2-self. One milestone theory, the L2 Motivational Self System (LMSS), was proposed by Dörnyei (2005) which explains the three facets of L2 motivation: Ideal L2 self, Ought-to L2 self and L2 learning experience.

The LMSS has been adopted as the theoretical framework in many current studies that explore L2 motivation characteristics and the effects of motivation on L2 learning. Yet most motivation studies focus on the overall language skills

rather than a specific domain of language. There are only a few motivation studies exploring how L2 motivation influences the development and learning of speaking, writing, listening and reading. Likewise, vocabulary learning motivation did not receive much attention in the pertinent literature. In addition, as L2 motivation explains learners' active and persistent learning, it should reflect on the learning process in particular the use of certain learning strategies. Studies on language learning strategies are mainly in line with the taxonomy by Oxford (1990) which includes four main types: cognitive, metacognitive, social and affective strategies. Yet how L2 motivation is related to these strategies remains underexplored. More recently, self-regulation has received increasing attention in SLA literature (e.g., Csizér & Tankó 2017), which is a closely related construct to L2 motivation and learning strategies. Thus, it is of great value to study the possible relationship between these variables.

Based on the reviewed scholarly literature, it can be seen that there is limited research on L2 vocabulary learning motivation even though it is well acclaimed as essential for successful L2 learning, especially in an acquisition-poor environment such as the EFL context in China. Moreover, how do L2 motivation reflects on the learning strategies and self-regulation still needs more explorations in the scholarship. Therefore, the present study aims to explore the characteristics of L2 vocabulary learning motivation and investigate how learning strategies and self-regulation are related to the L2 motivation.

Methods

The study employed a mixed-method approach with 47 college EFL students from Beijing, China, among whom 10 students had semi-structured interviews with the author. During the data collection procedure, two main instruments were used for quantitative data collection, namely a motivation questionnaire and a learning strategies questionnaire. The motivation questionnaire was adapted based on LMSS as well as the study by Taguchi et al. (2009). There are 5 scales with 42 items exploring the characteristics of L2 motivation: intrinsic motivation (8 items), instrumentality-promotion (11 items), instrumentality-prevention (8 items), ideal L2 self (8 items) and ought-to L2 self (10 items). These scales were tested in relevant studies (e.g., Taguchi et al. 2009) to be context-matched with the present study. The learning strategies questionnaire contains 28 items that tap into the main types of learning strategies such as cognitive strategies and metacognitive strategies. As for the qualitative data collection, the questions in the semi-structured interviews targeted the vocabulary learning motivation, learning strategies and, in particular, the self-regulation (e.g., What do you think of your willpower to achieve the vocabulary learning goals, if any, which you set for yourself?). Overall, the procedure contains the completion of a background questionnaire and two questionnaires, after that, ten of the participants were interviewed.

Results and discussion

The quantitative data were analysed using the principal component factor analysis, and Pearson correlation analysis; and thematic analysis was used to analyse the qualitative data.

The five constructs of motivation in vocabulary learning emerged in this research context with the overall Cronbach's alpha up to .869. Further explorations of these five constructs reveal that two components can be extracted from the instrumentality-promotion, namely promotion-general and promotion-study abroad. These two new factors were not significantly correlated (r=.093), suggesting that they are two separate constructs. The instrumentality is consistent with the literature that, learners from China as well as some other Asian countries, are likely to be instrumentally motivated in language learning. And because the participants in the study were senior college students who were graduating, a large number of them would choose to study abroad, the promotion-study abroad factor thus emerged as one motivational feature.

In this study, the Ideal L2 self was found to be the strongest factor for vocabulary learning whereas the Ought-to L2 self was the weakest one. The mean of these scales were significantly different (t (46)=12.014, p<.001), which indicates that, for these participants, they learn vocabulary because they want to be better and ideal selves in the future rather than being pushed to. The present study also confirmed the significant correlations between Ought-to L2 self and prevention (r=.663), and between Ideal L2 self and promotion-study abroad (r=.291), which validates the theoretical system of the LMSS by Dörnyei (2005). Besides, Ought-to L2 self was significantly correlated with the promotion-general, which is in line with the results in Taguchi et al. (2009) such that in China, young adults after college have to take care of their parents and families as their responsibilities, and so people would like to get promotion and secure higher income to support the families. Thus, Ought-to L2 self and promotion are correlated in this research context.

As regards the vocabulary learning strategies, the study shows that participants did not use many learning strategies to learn vocabulary, and overall, they employed significantly more cognitive strategies than the metacognitive strategies (t (46)=13.23, p<.001). The results also show that the metacognitive strategies are correlated with two promotion factors (e.g., general and study abroad). This result is expected as metacognitive involves more planning, reviewing in the learning process, the "promotion" will then trigger the use of such strategies.

The interview data were transcribed and performed with thematic analysis, and the five aspects (i.e., commitment control, metacognition, sanitation, emotion and time & environment) of self-regulation were examined. In general, these ten interviewees did not have a high level of self-regulation. More specifically, most of them believed that they did not have strong willpower to

achieve the vocabulary learning goals and may easily give up their plans and goals. And this is related to their procrastination in vocabulary learning, for example, they may have a plan to learn and review vocabulary but always procrastinate thereby failing to complete the learning goals. Also, these learners found that vocabulary learning was boring, even though some of them try to learn vocabulary in different ways such as using mobile applications. Besides, participants would eliminate the negative surroundings such as going to the study room in the library or other quiet places to learn vocabulary. Therefore, we can see that these participants had a low self-regulated capacity in learning vocabulary, they lack commitment control and metacognitive control; procrastination was a serious problem they have. Yet, they would control the sanitation and emotion in the learning process.

In conclusion, this study explored and confirmed the instrumentality of Chinese EFL learners in vocabulary learning and further revealed two components of instrumentality promotion. The motivation characteristics are related to learners' use of learning strategies and their self-regulated behaviours. Pedagogical implications can be drawn from the study that more attention should be paid to the vocabulary learning motivation and learning strategies of Chinese EFL learners, even though in the senior year, they still need to learn English vocabulary for some instrumental reasons, and some of them may have issues in this learning process, thus some guidance and instructions are necessary to be provided for students to facilitate the vocabulary learning.

References

Deci, E.L., Ryan, R.M. 1985. Intrinsic motivation and self-determination in human behavior. London: Plenum Press

Dörnyei, Z. 2005. The psychology of the language learner: Individual differences in second language acquisition. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z., Ushioda, E. (eds.) (2009). Motivation, Language Identity and the L2 Self. Bristol, UK: Multilingual Matters.

Gardner, R.C. 1985. Social psychology and second language learning: The role of attitudes and motivation. London: Edward Arnold Ltd.

Oxford, R.L. 1990. Language Learning Strategies: What Every Teacher Should Know. New York: Newbury House.

Taguchi, T., Magid, M., Papi, M. 2009. The L2 motivational self system among Japanese, Chinese and Iranian learners of English: A comparative study. Motivation, language identity and the L2 self, 36, 66-97.

Sentence repetition in German-speaking individuals with Down syndrome

Bernadette Witecy, Tatjana Tolkmit, Martina Penke
Department of Special Education and Rehabilitation, University of Cologne, Germany
https://doi.org/10.36505/ExLing-2020/11/0055/000470

Abstract

The present study investigated the morphosyntactic abilities of German-speaking individuals with Down syndrome (DS) employing a sentence repetition task. In addition, a nonword repetition task was used to assess verbal short-term memory. The performance of 16 children/adolescents with DS was compared to that of 10 typically developing (TD) children. Group comparisons as well as the inspection of standard scores that were determined based on nonverbal mental age indicated a significant morphosyntactic impairment in most individuals with DS that could neither be solely attributed to the general cognitive delay nor to the observed deficit in verbal verbal short-term memory. Further qualitative results are presented.

Keywords: down syndrome, developmental language disorder, morphosyntax, sentence repetition, verbal short-term memory

Introduction

Individuals with Down syndrome (DS), a neurodevelopmental disorder commonly causing intellectual disability, display noticeable problems in the comprehension and production of syntactic constructions and inflectional morphology (Abbeduto et al. 2007). Verbal short-term memory (VSTM) is another domain where individuals with DS have been found to be significantly impaired (Baddeley, Jarrold 2007) and is discussed as one possible source of the language difficulties in this group (Laws, Bishop 2004). The aim of the present study was to investigate the performance of German-speaking individuals with DS in a standardized sentence repetition (SR) task, commonly used to assess morphosyntactic abilities, and to compare it to that of typically developing (TD) children. Quantitative and qualitative analyses were performed to explore how the results are related to the performance in nonword repetition (NWR), a task that is suitable to assess VSTM in individuals with DS.

Method

Sixteen monolingual German-speaking children and adolescents with DS (9 male), aged 4;6 to 19;0 years (*M* 12;4 years), took part in the study. For two of them the parents reported a mild hearing loss. For the remaining participants with DS no permanent hearing loss had been diagnosed. Ten TD children were

included as controls. Their mean chronological age was 4;11 years (range: 3;11 to 5;7 years).

Nonverbal mental age (MA) in both groups was assessed using the SON-R 2.5-7 (Tellegen et al. 2007). It ranged from 4;0 to 6;5 years (*M* 5;1) in the individuals with DS and from 4;0 to 5;3 years (*M* 4;9) in the control group (no significant difference, $t(23) = .379$, $p = .708$).

VSTM was assessed using a NWR task consisting of 18 nonwords with increasing number of syllables and syllable complexity. The SR task included 15 sentences (length: 6-10 words), six of them semantically meaningful and nine with no meaningful content (e.g.: The clapping chairs sing a cup). Both tasks are part of a standardized test on German language development (SETK 3–5, Grimm et al. 2001). Raw scores (number of correctly repeated nonwords and correctly repeated words out of all sentences, respectively) and standard scores (T-scores), based on the MA of the participants, were calculated according to the test's manual. The maximum obtainable raw score was 18 for NWR and 119 for SR. In addition, results in SR were analysed qualitatively.

Results

The results of the two groups in NWR and SR are displayed in Table 1. Group comparisons revealed that the TD children performed significantly better than the DS participants in both tasks (comparison of raw scores: NWR: $t(24) = -2.319$, $p = .029$, SR: $t(24) = -6.891$, $p < .001$). Performance in NWR and SR was highly correlated in the DS group ($r(13) = .769$; $p = .001$), but not in the control group ($r(10) = .467$; $p = .173$).

In almost all participants with DS (13 of 16) standardized T-scores were higher in NWR than in SR. While eight of them performed according to MA expectations in NWR (T-score of 40 or above), only three did so in the SR task. Thus, five participants displayed a dissociation between performance in NWR and in SR.

A comparison of the number of correctly repeated content and function words in the SR task revealed that the participants with DS performed significantly worse on function words compared to content words ($t(15) = 8.101$, $p < .001$), whereas no such difference between word types could be found in the control group ($t(9) = .059$, $p = .954$) (see Table 1).

Errors in the SR task were classified as omissions, substitutions (including substitutions of the stem and/or the morphological marking) and additions. Only omissions and substitutions were further analysed. A repeated measures ANOVA with GROUP as between-subjects factor and WORD CLASS (function vs. content word) and ERROR TYPE (omission, substitution) as within-subject factors yielded a significant ERROR TYPE x GROUP interaction ($F(1,24) = 5.273$, $p = .031$). The DS group produced more omissions (62.9%) than substitutions (37.1%). The reverse held for the TD participants (44.0% vs. 56.0%). The three-way interaction between GROUP, WORD CLASS and ERROR TYPE was also

significant ($F(1,24) = 16.461$, $p < .001$). This was due to the higher number of omissions of function words in the DS group (see Figure 1). The other effects were not significant ($p > 0.5$).

Table 1. Mean scores for NWR and SR (standard deviations in parentheses). T-scores based on MA in the DS group.

	DS	TD
NWR (raw scores)	6 (10.4)	9.5 (3.6)
NWR (T-scores)	38.8 (10.4)	49 (15.5)
SR (raw scores)	43.7 (19.7)	93.2 (14.2)
SR (T-scores)	33.4 (7.1)	55.3 (7.1)
Content words (% correct)	47.7% (15.8)	80.1% (10.5)
Function words (% correct)	27.5% (20.5)	80.0% (11.0)

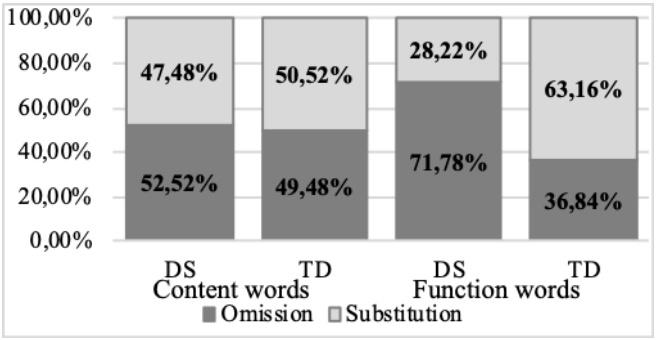


Figure 1. Percentage of omissions and substitutions for each group. Number of errors analyzed: DS: n = 948, TD: n = 258.

We furthermore examined the nature of the produced utterances in the SR task and graded them into four categories: (1) verbatim responses, (2) imitation with changes, but grammatical with respect to word order, (3) ungrammatical utterances, and (4) one-word-utterances. Whereas the TD children mainly produced grammatical sentences (89.7%), in the DS group almost two third of the utterances (64.8%) were ungrammatical (Table 2).

Table 2. Results of the syntactic analysis (Number of analysed utterances: DS: n = 227, TD: n = 175).

	verbatim	correct word order	incorrect word order	single-word
DS	2.6%	32.6%	59.0%	5.7%
TD	40.0%	49.7%	10.03%	0.0%

Discussion

The findings indicate a severe morphosyntactic impairment in most of the participants with DS that cannot simply be attributed to the general cognitive delay as most participants performed below MA expectations. In addition, even though results in NWR and SR were correlated in the DS group, the observation that performance in NWR was generally better and often dissociated from performance in SR suggests that the morphosyntactic impairment cannot be solely attributed to a deficit in VSTM. Finally, the higher error-proneness of function words, the observation that function words were omitted rather than substituted and the larger number of ungrammatical utterances in the DS group all point to a grammatical component of the disorder. Our findings support the assumption that morphosyntactic deficits in individuals with DS cannot be reduced to limitations in VSTM (Penke & Wimmer 2020).

Acknowledgements

We would like to thank all participants and their families as well as everyone who was involved in data collection and analysis. This research was supported by the German Science Foundation (DFG, grant PE 683/3-1).

References

Abbeduto, L., Warren, S., Conners, F. 2007. Language development in Down syndrome. Mental Retardation and Developmental Disabilities Research Reviews 13 (3), 247–261.

Baddeley, A., Jarrold, C. 2007. Working memory and Down syndrome. Journal of Intellectual Disability Research 51 (12), 925–931.

Grimm, H., Aktas, M., Frevert, S. 2001. SETK 3–5: Sprachentwicklungstest für drei- bis fünfjährige Kinder. Göttingen, Hogrefe.

Laws, G., Bishop, D. 2004. Verbal deficits in Down syndrome and specific language impairment. International Journal of Language & Communication Disorders 39 (4), 423-451.

Penke, M., Wimmer, E. 2020. Verbal short-term memory and sentence comprehension in German children and adolescents with Down syndrome – beware of the task. First Language 40 (4), 367-389.

Tellegen, P., Laros, J., Petermann, F. 2007. SON-R 2,5-7. Non-verbaler Intelligenztest. Göttingen, Hogrefe.

L1 Russian L2 Chinese speakers' acquisition of the Chinese conjunction *hé*

Xin Yan, Lemei Peng, Shanshan Yan
School of Chinese as a Second Language, Peking University, China
https://doi.org/10.36505/ExLing-2020/11/0056/000471

Abstract

From the perspective of features, this paper conducts an empirical study to examine L1 Russian L2 Chinese Speakers' acquisition of the Chinese conjunction *hé*. Specifically, the result shows that the vitality of an extraneous feature decreases until learners' language proficiency reaches an intermediate level. Eventually, it is difficult for learners to remove this extraneous feature. It suggests that this feature becomes dormant.

Keywords: SLA, Russian-speaking L2 learners, Chinese conjunction *hé*, feature

Introduction

The Chinese conjunction *hé* is one of the most common functional words in Mandarin Chinese, and its core function is to connect two Determiner Phrases (DPs) that indicate a coordinating relation (Lü and Li, 1980).

When L1-Russian speakers learn this word, they always match it with its Russian equivalent - the conjunction *u*. However, the function of *u* is not exactly the same as *hé* – it also indicates a coordinating relation and can connect DPs, but it can also connect Complementizer Phrases (CPs) (Кузнецов, 1998).

This research marks the feature that represents the coordinating relation as [COOR], the feature that can connect DPs as [DP-LINK], and the feature that can connect CPs as [CP-LINK]. The feature sets as shown in Figure 1.

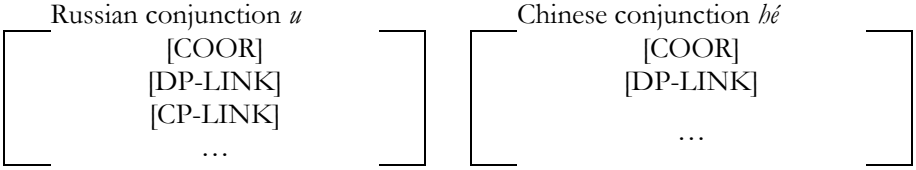


Figure 1. The feature sets.

Lately, *feature* has become a core concept under the framework of *the Minimalist Program* (Chomsky, 1995). The Feature Reassembly Hypothesis (Lardiere, 2008, 2009) explained that if the feature sets of the L1 and L2 are similar but not identical, then the learner will map the features of the L1 to the L2 lexical item, and features will be reassembled if successful. Furthermore, the Dormant Feature Hypothesis (Yuan, 2014) claims that there is not always enough evidence to help learners confirm or disconfirm features in the L2 input, and these features gradually lose their vitality and remain in a dormant state.

Empirical study

Research questions

a. Can learners successfully establish a mapping between *hé* and *u*?
b. Can learners map the [COOR] and [DP-LINK] onto the L2 lexical item?
c. Can learners map the [CP-LINK] onto the L2 lexical item? If they can, will it be removed as the learner's language proficiency improves?

Participants

There were 18 L1 Chinese speakers and 66 L1 Russian L2 Chinese speakers. L2 speakers were divided into 3 proficiency groups.

Instruments and procedures

Translation task

This task required the participants (only L2 groups) to translate a Russian sentence with the conjunction *u* into Chinese with 1 point for a correct answer and 0 points for an incorrect answer. When translating "DP_1 *u* DP_2", the participants should translate *u* into *hé*, and when translating "CP_1, *u* CP_2", the participants should not translate *u* into *hé* because *hé* cannot connect CPs.

Acceptability judgment task

This task required the participants to judge the acceptability of five types of sentences, including: "DP_1 *hé* DP_2"; "CP_1, *hé* CP_2"; "CP_1, CP_2"; "CP_1"; "CP_2". The score depended on the participants' acceptability.

Sentence-making task

This task required the participants to connect the given words into grammatical sentences with 1 point for a correct answer and 0 points for an incorrect answer. There were two types of sentences: "DP_1 *hé* DP_2" and "CP_1, *hé* CP_2".

Results

Translation task

Table 1. The mean scores of the translation task.

Groups	Number	DP_1 *hé* DP_2	CP_1, *hé* CP_2
elementary	19	0.92	0.59
intermediate	27	0.92	0.80
advanced	20	0.98	0.84
native	18	N/A	N/A

In type "DP_1 *hé* DP_2", one-way ANOVA analysis indicated that there were no significant differences between the groups (F (2, 261) = 1.485, p>0.05). However, in type "CP_1, *hé* CP_2", there were significant differences in each group (F (2, 261) = 7.643, p<0.05), while the *post hoc* Scheffé tests indicated that

the intermediate and advanced groups tended to reject using *hé* to translate the Russian coordinating compound sentence with *u* (p<0.05).

Acceptability judgment task

Table 2. The mean scores of the acceptability judgment task.

Groups	Number	DP_1 hé DP_2	CP_1, hé CP_2	CP_1, CP_2	CP_1	CP_2
elementary	19	3.66	2.87***	3.65	3.79 *	3.73 *
intermediate	27	3.66	1.98***	3.76	3.89	3.81
advanced	20	3.90	1.91***	3.89	3.96	3.95
native	18	3.75	1.17	3.79	3.97	3.94

Notes. *= significantly different from the NS Group at $p < 0.05$; *** = significantly different from the NS Group at $p < 0.001$.

One-way ANOVA analysis shows that there were no significant differences between the groups in type "DP_1 *hé* DP_2" (F (3, 332) = 3.432, p>0.05) and type "CP_1, CP_2" (F (3, 332) = 2.689, p<0.05). However, in type "CP_1, *hé* CP_2", a very significant difference was found (F (3, 332) = 38.160, p<0.001), while the *post hoc* Scheffé tests indicated that there were very significant differences between L1 and L2 groups (p<0.001). Moreover, the elementary group had a very significant difference from the intermediate and advanced groups (p<0.001). A paired sample T-test analysis showed that there were very significant differences between type "CP_1, CP_2" and "CP_1, *hé* CP_2" of each L2 groups (t (72) = -30.528, p<0.001).

Sentence-making task

Table 3. The mean scores of the sentence-making task.

Groups	Number	DP_1 *hé* DP_2	CP_1, *hé* CP_2
elementary	19	0.87	0.34 ***
intermediate	27	0.92	0.69 ***
advanced	20	0.96	0.63 ***
native	18	0.94	1.00

Notes. *** = significantly different from the NS Group at $p < 0.001$.

One-way ANOVA analysis shows that there were no significant differences in type "DP_1 *hé* DP_2" (F (3, 332) = 1.822, p>0.05), whereas in type "CP_1, *hé* CP_2" a very significant difference was found (F (3, 332) = 30.252, p<0.001), the *post hoc* Scheffé tests indicated that there were very significant differences between L1 and L2 groups (p<0.001).

Discussion

In the translation task, the L2 groups could all translate *u* that connects DPs to *hé*. This shows that in the initial stage, learners could match *hé* with *u*, and map the [COOR] and [DP-LINK] to the L2 lexical item. However, the mean scores of the elementary group was significantly lower than that of other two learner's

groups, which indicates the [CP-LINK] feature was attached to the elementary learners' L2 grammars.

In the acceptability judgment task, the mean score of all L2 groups of "CP_1, *hé* CP_2" was significantly different from the L1 group's, the [CP-LINK] feature was active in the learners' L2 lexical item. In addition, the elementary group had a very significant difference from the intermediate and advanced groups, indicating that the vitality of the [CP-LINK] feature decreases with the improvement of language proficiency, but the feature's vitality did not change after language proficiency reached a certain level. What's more, learners were sensitive to the difference between "CP_1, *hé* CP_2" and "CP_1, CP_2", but all learners could not reject "CP_1, *hé* CP_2", which means that the [CP-LINK] feature still randomly affects the judgments of these learners. According to the Dormant Feature Hypothesis (Yuan, 2014), it is believed that that the [CP-LINK] feature has entered a dormant state.

In the sentence-making task, The L2 groups could produce a sentence using *hé* to connect DPs, just like the L1 group. However, even advanced learners still produce type "CP1, *hé* CP2" sentences, which means that even advanced learners could not completely remove the [CP-LINK] feature from their L2 lexical item.

Conclusion

This study investigated Russian speakers' acquisition of the Chinese conjunction *hé* from the perspective of features. It found that at the beginning of learning, learners matched *hé* with the Russian conjunction *u*. Also, learners mapped the features of *u* [COOR, DP-LINK, CP-LINK] onto the L2 lexical item *hé*. As language proficiency improved, the vitality of the [CP-LINK] feature decreases, but the feature still randomly affects the behaviors of advance learners. It is believed that due to the lack of negative evidence, the [CP-LINK] feature has entered a dormant state.

References

Chomsky, N. 1995. The Minimalist Program. Cambridge, Mass, The MIT Press.

Lardiere, D. 2008. Feature assembly in second language acquisition. In Liceras JM, Zobl H, and Goodluck H (eds) The role of formal features in second language acquisition, 106–40, Mahwah, NJ, Lawrence Erlbaum.

Lardiere, D. 2009. Some thoughts on the contrastive analysis of features in second language acquisition. Second Language Research, 25(2), 173-227.

Lü, S., Li, L. 1980. Xiandai hanyu babai ci [Eight Hundred Words of Modern Chinese]. Beijing, The Commercial Press.

Yuan, B. 2014. 'wh-on-earth' in Chinese speakers' L2 English: Evidence of dormant features. Second Language Research, 30(4), 515-549.

Кузнецов, С. А. 1998. Большой толковый словарь русского языка [Great Dictionary of Russian language]. St. Petersburg: Norint, 1998.

Individual differences in Mandarin focus production

Yike Yang[1], Si Chen[1,2]

[1]Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong SAR

[2]The Hong Kong Polytechnic University-Peking University Research Centre on Chinese Linguistics, Hong Kong SAR

https://doi.org/10.36505/ExLing-2020/11/0057/000472

Abstract

This paper investigated whether and how individual speakers of Mandarin Chinese (Mandarin) mark prosodic focus (broad focus vs verb focus) differently in their production, and tested focus effects on mean F0, duration and intensity. The findings indicated the role of the three acoustic cues in Mandarin focus marking at both the group and individual levels. Meanwhile, the individual data showed great variations among speakers in terms of the extent to which the cues were employed. It is proposed that the dynamics of acoustic cues should be considered in future studies and caution should be taken when selecting stimuli for focus perception studies.

Keywords: individual differences, speech production, prosody, focus, Mandarin

Introduction

Research into individual speaker differences in speech production is of great significance for our understanding of the way speakers signal linguistic contrasts (Smith & Hawkins 2012). However, little is known about how individual speakers vary in using prosodic cues to mark information structure (Ouyang & Kaiser 2015). Cross-linguistically, various acoustic cues are employed to signal information structure (Lee 2015). In Mandarin Chinese (Mandarin), focus influences the mean F0, duration and intensity in both local and global regions, wherein the on-focus region receives the greatest prominence and the post-focus region receives the least (Yang & Chen 2020).

This project aimed to investigate whether and how individual speakers of Mandarin mark prosodic focus differently in their production. Specifically, this study examined the individual differences in the implementation of mean F0, duration and intensity in pre-, on- and post-focus regions in the production of Mandarin broad focus and verb focus.

Methodology

Eleven native speakers of Mandarin (six females; aged: 24.72 ± 4.39) attended a production experiment at the Speech and Language Sciences Laboratory of the

Hong Kong Polytechnic University. All participants gave their written informed consent prior to the recording sessions.

Table 1. Focus types and question-answer pairs (focused regions underlined).

Focus types	Precursor questions	Answers
Broad focus	ni shuo shen me? 'What did you say?'	na wei yisheng he kafei. 'The doctor drinks coffee.'
Verb focus	na wei yisheng zenme kafei? 'What does the doctor do to the coffee?'	na wei yisheng <u>he</u> kafei. 'The doctor <u>drinks</u> coffee.'

The stimuli consisted of six subject–verb–object (SVO) declarative sentences with exactly the same structure. To make the production data more natural, the broad focus and verb focus utterances were elicited with precursor questions, as illustrated in Table 1. Only the answers were further processed. Relevant acoustic values, including mean F0, duration and intensity of each syllable, were extracted and analysed with linear mixed-effects modelling for each region (pre-, on- and post-focus regions, each corresponding to subject, verb and object positions, respectively).

Results

The group data are plotted in Figure 1, which revealed very clear focus effects. Linear mixed-effects models suggested that focus significantly increased mean F0 ($p < .001$), duration ($p < .001$) and intensity ($p = .005$) in the on-focus region. There was also a main effect of focus on F0 in the pre-focus region and a main effect of focus on F0 and intensity in the post-focus region (ps < .001).

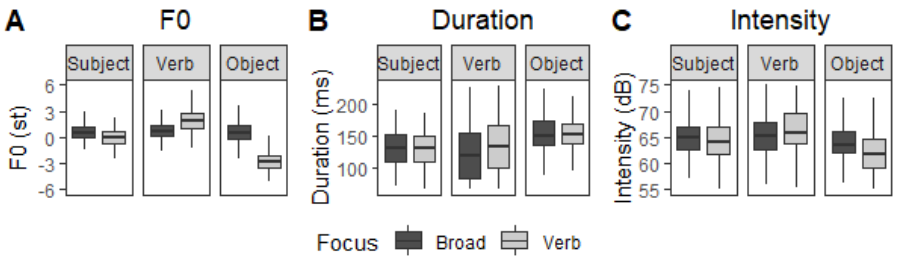


Figure 1. Group data of focus effects.

'Speaker' was then included as a fixed effect to examine whether individual speakers used the acoustic cues differently. Of the nine models (3 regions * 3 acoustic cues), eight showed a main effect of speaker (ps < .001). The only non-significant model was the one for on-focus F0, which reflected the fact that most speakers increased F0 when the constituent was under focus.

Next, individual models were fitted for each acoustic cue produced by each speaker in each region, and the results indicated huge variations among speakers in their use of the acoustic cues, as presented in Figure 2. For example, Speaker 1 did not use F0 (p = .715) or intensity (p = .890) to mark verb focus in the on-focus region, but there was a clear decrease of F0 and intensity in the pre- and post-focus regions (ps < .039; Figures (2A) and (2C)), which made the pronunciation of the verb prominent. Also, while most speakers increased on-focus F0, the degree of the increase can be as large as 3.519 ± 0.396 st for Speaker 2 (p < .001; Figure (2D)) and as small as 0.445 ± 0.128 st for Speaker 10 (p = .002; Figure (2G)). Besides, as shown in Figure (2E), Speaker 2 appeared to have lengthened duration in all the regions to mark verb focus.

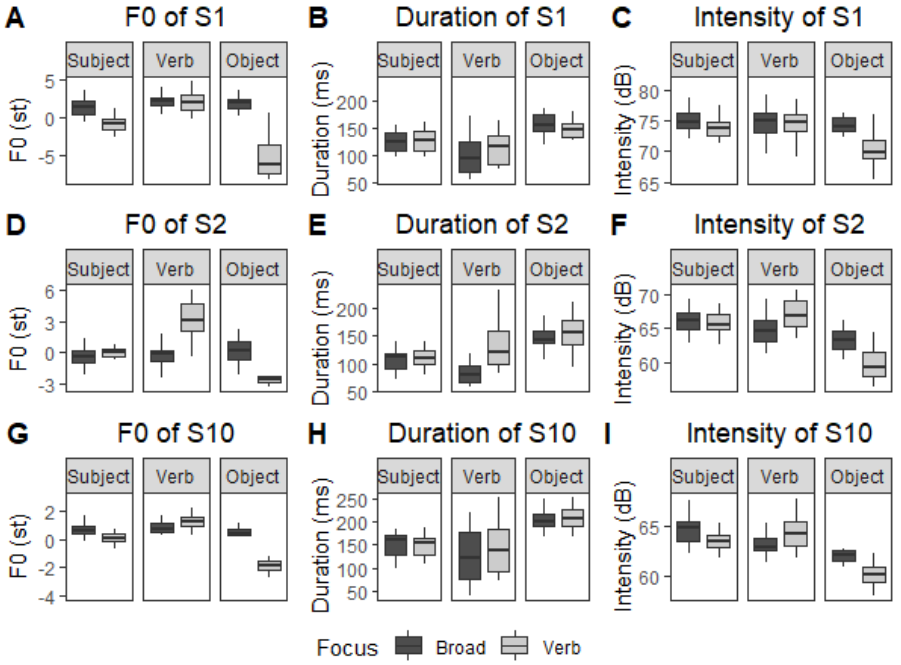


Figure 2. Individual data of focus effects ('S' for 'speaker').

Discussion and conclusions

This paper investigated how individual speakers differed in marking Mandarin broad focus and verb focus. At the group level, focus affected mean F0 in all the three regions and affected intensity in both on- and post-focus regions, but such effect for duration was observed in the on-focus position only. At the individual level, although every speaker made use of the three tested acoustic cues, there were great variations among speakers in terms of the extent to which the cues were employed.

Note that we converted the F0 values of each speaker from Hz to st individually, with the mean F0 of each speaker as the reference. Despite the fact that normalised scores were used, the F0 values exhibited large variations even under the broad focus condition. If we observe the mean F0 values from the subject position to the object position under broad focus, while Speaker 2 had a gradual increase (Figure (2D)), Speaker 10 showed an increase from subject to verb and then a decrease from verb to object (Figure (2G)). Also, the mean F0 values of Speaker 2 centred around 0 st (Figure (2D)) and the mean F0 values of Speakers 1 and 10 were well above 0 st (Figures (2A) and (2G)). The verb focus condition, needless to say, presented even more complex patterns, which require further examination. In addition to mean values, it is necessary to consider the dynamics of F0 and other acoustic cues to account for the observed individual differences in Mandarin prosodic focus (Roessig et al 2019).

Although the three acoustic cues were used by all the speakers, not every speaker employed these cues in each syntactic position. From our preliminary data, we postulate a possible hierarchy of the cues in marking Mandarin focus: F0 > intensity > duration. There still remain some issues for focus perception studies. First, given the different weights the cues carry in focus production, a follow-up question would be how much each acoustic cue contributes to Mandarin focus perception. Second, having observed very complicated patterns in focus production, one would ask whether there are also interactions among the cues in focus perception. Third, if speaker variation is the norm in focus production, we must be more cautious when selecting stimuli for focus perception tasks.

In conclusion, the findings indicated the role of mean F0, duration and intensity in Mandarin focus marking at both the group and individual levels. The dynamics of acoustic cues should be considered in future studies to account for the observed differences. Besides, issues for focus perception studies are also proposed.

References

Lee, Y.-C. 2015. Prosodic focus within and across languages. Unpublished PhD dissertation, University of Pennsylvania.

Ouyang, I. C., Kaiser, E. 2015. Individual differences in the prosodic encoding of informativity. In Fuchs, S. et al (eds.), Individual Differences in Speech Production and Perception, 147-188. Bern, Peter Lang.

Roessig, S., Mücke, D., Grice, M. 2019. The dynamics of intonation: Categorical and continuous variation in an attractor-based model. PLoS ONE, 14(5), e0216859.

Smith, R., Hawkins, S. 2012. Production and perception of speaker-specific phonetic detail at word boundaries. Journal of Phonetics, 40(2), 213-233.

Yang, Y., Chen, S. 2020. Revisiting focus production in Mandarin Chinese: Some preliminary findings. In Proc. Speech Prosody 2020, 260-264, Tokyo, Japan.

Quantity inferences in conditionals: a pilot experimental study

Natalia Zevakhina, Veronika Prigorkina
International Laboratory for Logic, Linguistics and Formal Philosophy, National Research University Higher School of Economics
https://doi.org/10.36505/ExLing-2020/11/0058/000473

Abstract

Relying upon Fillenbaum (1975), the paper reports on a more systematic experimental study of the role of the following factors (and levels) in the derivation of Conditional Perfection: negation (no negation, negation in an antecedent, negation in a consequent, negation both in an antecedent and a consequent), order of two clauses of a conditional sentence (if p, q; q, if p), and face vs. non-face speech acts (promises, threats vs. causals, temporals and contingent universals).

Keywords: quantity inference, Conditional Perfection, conditional, negation, face

Introduction

Quantity inferences which are derived on the basis of conditional sentences were coined as Conditional Perfection in Geis and Zwicky (1971), see the famous example: *If you mow the lawn, I'll give you five dollars*; inference: If you don't mow the lawn, I won't give you five dollars. Schematically, the example (1) has the structure 'if p, q' (where p and q are clauses) and the inference has the structure 'if not p, not q'. Relying upon Fillenbaum (1975), we conducted a more systematic study of the role of various types of speech acts and the role of negation in the derivation of Conditional Perfection. Additionally, we tested the role of clause order.

The first question (**Question A**) is whether the derivation of an inference from a negated antecedent and/or a negated consequent would take a time interval comparable to the processing of a conditional sentence with no negation. The prediction (**Prediction A**) is that conditional sentences with no negation yield inferences faster than conditional sentences with single negation (that is negation in an antecedent or in a consequent), which give rise to inferences faster than conditional sentences with double negation (that is negation in both an antecedent and a consequent). In other words, the prediction is as follows: if p, q >> if not p, q; if p, not q >> if not p, not q. The reason for the prediction lies in that negation is time-consuming. The second question (**Question B**) is whether the processing of an order antecedent + consequent would take a time interval comparable to the processing of an order consequent + antecedent. The prediction (**Prediction B**) is that the derivation of inferences from conditional sentences with the direct order (that is

antecedent + consequent) takes a lesser timing than the derivation of inferences from conditional sentences with the inverse order (that is consequent + antecedent). In other words, the prediction is as follows: if p, q >> q, if p. The reason for this is that, intuitively, the inference derivation starts from an antecedent (that is has a form "if p, q") and in case of "q, if p" the hearer firstly changes the order of the clauses and secondly derives an inference, and this presumably takes extra time. The third question (**Question C**) was whether face speech acts facilitate the derivation of Conditional Perfection since they concern damage to or respecting the hearer's interests. The prediction (**Prediction C**) is that speech acts that involve face give rise to more inferences and take lesser timing than speech acts that lack face.

Methods

We distinguished among five types of speech acts: promises and threats (face), causal and temporal sentences (no face), contingent universals (no face). Each of the five types was constructed as statements with no negation, single or double negation presented in the direct or inverse order. Moreover, according to the research questions A, B, C and predictions A, B, C, each of the five groups came into the following schematic varieties: 1) statements with no negation, with single negation, with double negation presented in the direct order: If p, q; If not p, q; If p, not q; If not p, not q; 2) statements with no negation, with single negation, with double negation presented in the indirect order: q, if p; q, if not p; not q, if p; not q, if not p.

Lexical contents of all sentences were different. Each of the schemas was represented in 2 versions distributed between 2 experimental lists. We did this to decrease a potential effect that some sentence might have had. We generated speech acts x order x negation x 2 versions for each combination distributed between 2 lists, with 40 stimuli per list, see (1)–(4). We used an inference task, cf. Geurts and Pouscoulous (2009) a.o. Participants were presented with a conditional sentence on a slide followed (on the next slide) with the question whether it is possible to make an inference (Conditional Perfection) as well as with yes (key *G*) and *no* (key *J*) answers. There was a time interval of 10 sec to read a conditional sentence and a time interval of 15 sec to answer the question. The latter interval is maximum, that is, participants had to choose an answer within this interval and their reaction times were recorded. The experiment was conducted in Russian.

(1) Esli ty pol'još mne cvety, ja nakormlju tebja pirogami. Možete li Vy sdelat' iz etogo vyvod, čto esli ty ne pol'yoš mne cvety, ja ne nakormlju tebja pirogami?
'If you water my flowers, I will give you the pies. Would you infer from that that if you don't water the flowers, I won't give you the pies?' (Promise; If p, q)

(2) Esli Anja ne nadenet šarf, ona zamjorznet na ulice. Možete li Vy sdelat' iz etogo vyvod, čto esli Anja nadenet šarf, ona ne zamjorznet na ulitse?
'If Anja does not put on her scarf, she will get freezed outside. Would you infer from this that if Anja puts on her scarf, she won't get freezed outside?' (Causal; if not p, q)

(3) Borja ne prinimaet zvonki s raboty, esli on otdyxaet na more. Možete li Vy sdelat' iz etogo vyvod, čto Borja prinimaet zvonki s raboty, esli on ne otdyxaet na more?
'Borja does not answer phone calls if he is having a sea vacation. Would you infer from this that Borja answers phone calls if he is not having a sea vacation?' (Temporal; not q, if p)

(4) Desertnoj vilkoj ne pol'zujutsja, esli v menju ne predusmotren desert. Možete li Vy sdelat' iz etogo vyvod, čto desertnoj vilkoj pol'zujutsja, esli v menju predusmotren desert?
'One does not use a dessert fork if the menu does not contain a dessert. Would you infer from this that one uses a dessert fork if the menu contains a dessert?' (Contingent Universal; not q, if not p)

46 fillers were statements which denote sequences of events and were followed with a question that violates the order of the events. They were expected to receive a definite "yes" response or a definite "no" response, see (5)-(6). 6 of them were presented at the beginning of the experiment in order to train participants for the further experimental materials. The same fillers were used in both lists.

(5) Denis polil cvety, prežde čem oni zavjali ot nedostatka vody. Možete li Vy sdelat' iz etogo vyvod, čto Denis polil cvety, a potom oni zavjali ot nedostatka vlagi?
'Denis watered the flowers before they withered because of the lack of water. Would you infer from this that Denis watered the flowers and then they withered from the lack of water?' (Expected answer is "no").

(6) Eva vyšla iz kinozala prežde, čem kino zakončilos'. Možete li Vy sdelat' iz etogo vyvod, čto Eva vyšla iz kinozala, a potom kino zakončilos'?
'Eva left the cinema before the movie ended. Would you infer from this that Eva went out from the cinema and then the movie ended?' (Expected answer is "yes").

Due to a big number of stimuli and fillers, each 20 items were followed with a small break for 10 sec in order to give some rest for participants. The

experiment was conducted via IbexFarm, which is a free online platform for conducting experiments. There were 68 participants (49 female, age range=19–39, mean age=22).

Results

Out of 2720 received responses, 80 responses were removed due to null answers and 238 responses were excluded due to the extraordinarily slow or fast RTs (>10000 ms and <250 ms respectively). To analyze the results of the study, generalized mixed-effects model for yes/no-answers and linear mixed-effects model for RTs were used. Focusing on the stimuli only, the overall acceptance of Conditional Perfection was significantly high, with average 76.34% ($p<.001$). Both for answers and RTs, the interaction between speech acts vs. negation was significant ($p<.05$), but not between negation vs. order or speech acts vs. order ($p>.05$). Pairwise comparisons among types of negation for answers and RTs revealed that double negation significantly facilitates the derivation of Conditional Perfection and is processed faster than the single negation or no negation ($p<.01$). This suggests that Prediction A was not fulfilled. One plausible reason for this is that in case of double negation, both clauses of a conditional are marked with negation and, in this sense, are parallel in processing. Another plausible reason is that if a sentence contains double negation, its quantity inference includes no negation, that is double negation is tantamount to affirmation (cf. $\neg\neg p=p$). Pairwise comparisons among the types of clause order for answers and RTs revealed no significant difference between direct vs. inverse types of order ($p>.05$). Prediction B was not fulfilled. This means that the order factor does not seem to be relevant in the derivation of quantity inferences in conditionals. Finally, pairwise comparisons among speech acts for answers and RTs revealed that face speech acts facilitate the derivation of Perfection. Prediction C was fulfilled. This suggests that inferential reasoning is dependent upon the face factor.

References

Fillenbaum, S. 1975. If: Some Uses. Psychological Research 37, 245-260.

Geis, M.L., Zwicky, A.M.. 1971. On invited inferences. Linguistic inquiry 2, 561-566.

Geurts, B., Pouscoulous, N. 2009. Embedded implicatures?!? Semantics and pragmatics 2, 1-34.

Index of names

ExLing 2020
Ebook ISSN: 2529-1092
Ebook ISBN: 978-618-84585-1-2