ExLing

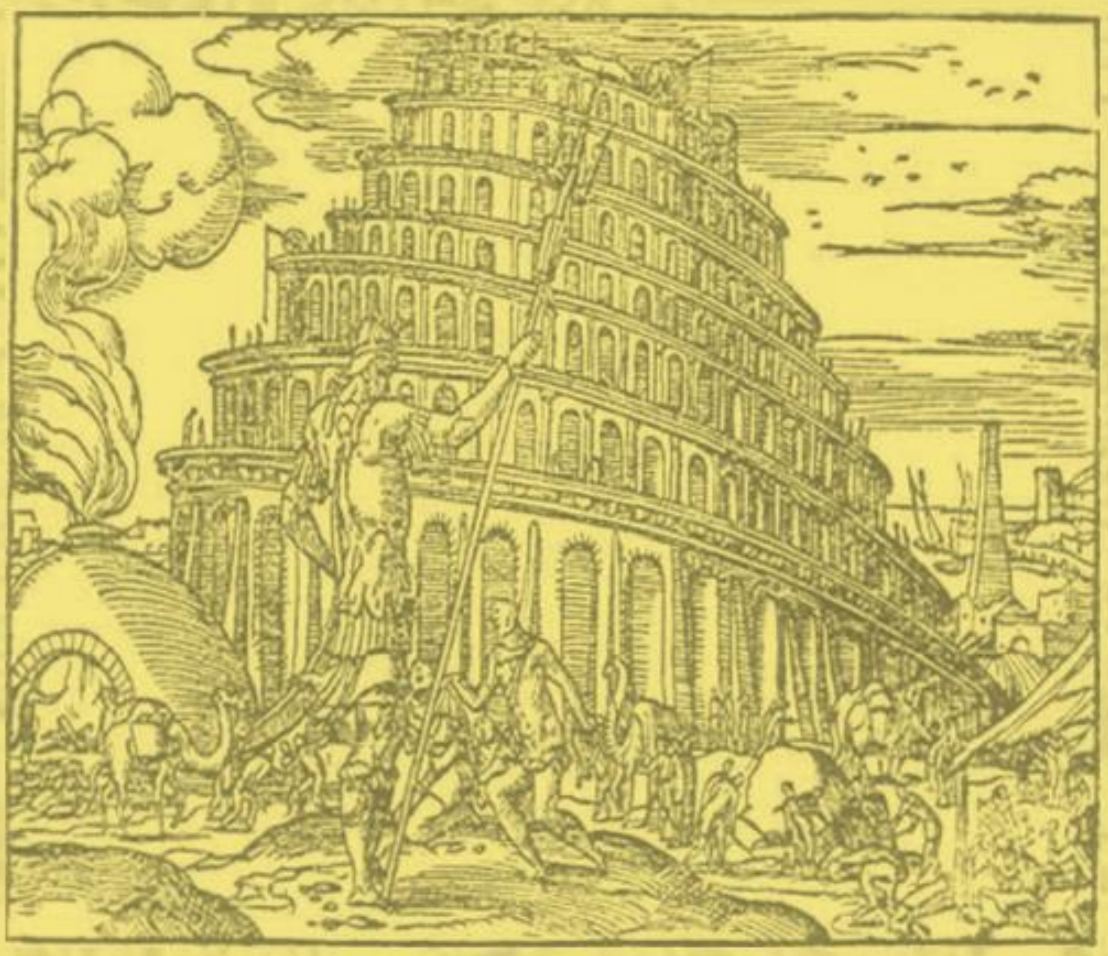
International Society of Experimental Linguistics

ExLing 2023

Proceedings of 14th International Conference of Experimental Linguistics

18-20 October 2023
Athens, Greece

Book of abstracts



HELLENIC REPUBLIC
National and Kapodistrian University of Athens
EST. 1837

International Society of Experimental Linguistics

ExLing 2023

Proceedings of 13th International Conference of Experimental Linguistics

18-20 October 2023

Athens, Greece

Book of abstracts

ExLing 2023
Proceedings of 14th International Conference of Experimental Linguistics

Published by ExLing Society
Electronic edition
ExLing 2023
Athens, Greece
ISSN: 2529-1092
ISBN:
DOI: 10.36505/ExLing-2023/14

Foreword

Welcome to the 14th International Conference of Experimental Linguistics ExLing 2023 Athens. This year, our hybrid set up enables us to meet in-person thanks to the hospitality of the University of Athens, enjoying at the same time the excitements of the Athenian life and the Attica weather.

As an international forum for generations of linguists, our Society is collectively devoted to analyzing language through the development of experimental methods in Linguistics. We encourage new and established researchers to participate in and discuss developments in linguistic research and related experimental methodologies.

It all began in 2006 in Athens, with the first ExLing Conference (at the time an International Speech Communication Association (ISCA) ExLing Workshop). Subsequently, it was hosted in cities such as Paris, Saint Petersburg, Heraklion, Lisbon and has been established since 2019 as the annual International Conference of Experimental Linguistics.

ExLing 2023 has come back to Athens and we are excited to return to the roots of the ExLing conference once more, this time with the added benefit of virtual conference attendance which opens new horizons and allows us to reach a wider audience.

This volume includes the proceedings of ExLing 2023 Athens. In addition to the main conference, ExLing 2023 Athens is hosting three special sessions encompassing diverse and energetic research domains focusing on experimental methods and current topics in Experimental Phonetics, Language Education and Language Pathology.

The launching of the Language Application Exhibition presents a new aspect of the ExLing 2023 Athens conference, which is expected to unify diverse language applications and promote interdisciplinary advances in major areas of technological language aids.

We would like to thank all ExLing 2023 Athens participants and our keynote speakers Edward Flemming, Gareth Gaskell, Gaetanelle Gilquin, as well as colleagues from the International Advisory Committee, and the Review and Organization Committees for their contribution to the successful outcome of the Conference.

Antonis Botinis
ExLing Society

Contents

Keynote papers

Research papers

Generating and parsing f0 contours using a functional model of f0 production

Keynote paper
Edward Flemming
MIT, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

I will present ongoing work aimed at developing a framework for formulating phonetic grammars of tone realization that can derive complete f0 contours from phonological specifications.

The proposed framework consists of two main components: A set of weighted constraints that select and locate optimal tone targets for a given phonological representation, and a functional model of f0 production that maps time-aligned tone targets to f0 contours (based on Birkholz & Xu 2018). These two components work in parallel: the constraints that determine tone targets evaluate properties of the resulting f0 contour.

In addition to generating f0 contours, the model of f0 production can also be fitted to observed f0 contours to infer the underlying f0 targets, effectively parsing the f0 contour (cf. Öhman & Lindqvist 1965, Birkholz & Xu 2018). This proves particularly valuable when analyzing tones with targets that are difficult to locate because they do not correspond to a local maximum or minimum, such as tone targets that result in a transition from a falling f0 movement to a relatively level plateau, or vice versa. These kinds of f0 events are often referred to as 'elbows',

The framework will be illustrated through analyses of two tonal phenomena that involve elbow targets: (i) the timing of the L- phrase accent in English H*L- intonational melodies, and (ii) realization of the Mandarin rising tone.

References

Birkholz, P., Xu, Y. 2018. Estimation of pitch targets from speech signals by joint regularized optimization. 26th European Signal Processing Conference, 2089-2093.

Öhman, S. Lindqvist, J. 1965. Analysis-by-synthesis of prosodic pitch contours. STL-QPSR 6(4), 1-6.

Sleep to improve your language

Keynote paper
Gareth Gaskell
University of York, UK

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Throughout our lives as language users we need to adapt to new linguistic circumstances and refine our accumulated knowledge. In this talk I will discuss the role that sleep plays in this plasticity. I will describe several situations in which new language learning becomes better integrated with existing knowledge following a period of sleep. For example, when we encounter an unfamiliar word, its ability to engage in lexical processing is strengthened by sleep. In particular, new words are better able to influence the speed of recognition of their existing neighbours after sleep (enhanced lexical competition). Similarly, the acquisition of new phonotactic constraints in speech production show a stronger influence on speech errors after sleep. So far as we know, this profile of learning operates through development and adulthood, and applies to both first and second language acquisition.

I will discuss how a complementary systems model applied to sleep can accommodate these findings. Key components of this model support the rapid episodic encoding of new linguistic material via the hippocampus, plus a more stable long-term repository of language knowledge in cortical regions. Sleep provides one means by which dialogue between these components can update more crystallised knowledge, with specific aspects of sleep (sleep spindles and slow oscillations) implicated in the process.

Although this model applies most obviously to situations where we have something novel to learn, such as an unfamiliar word, recent evidence is emerging that the model applies more broadly. Even when linguistic materials are fully familiar there is learning to be done, in terms of updating our semantic knowledge associated with words and keeping track of conversations and texts. Therefore, I will also describe recent evidence suggesting that the same encoding/consolidation system can operate to support plasticity and learning in these domains.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Experimental methodologies and corpus linguistics in language learning research

Keynote paper
Gaetanelle Gilquin
Université Catholique de Louvain, Belgium

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

No single methodology can provide a full picture of what language learning looks like. In this presentation, I will argue for the combination of experimental methodologies and corpus linguistics in language learning research. I will show how this combination makes it possible to take advantage of the strengths of each approach while compensating for their weaknesses. I will also highlight the challenges involved in combining experimental methodologies and corpus linguistics and will provide some suggestions as to how these challenges can be overcome.

The talk will be illustrated with case studies relying on different experimental methodologies and corpus linguistic techniques. One case study will centre around language learners' knowledge of collocations with make. The results of fill-in-the-blank exercises and acceptability judgments will be compared with an error analysis and a collostructional analysis of the corpus data. The role of L1 influence will be investigated in both types of data.

The other case study will deal with language learners' writing processes. It will be based on textual data combined with keylogging, eye-tracking and stimulated recall. The contribution of each type of data will be examined and it will be shown how the tools and techniques of corpus linguistics can help with the analysis. Some findings will be presented, including findings from data produced by language learners with dyslexia.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Pronoun resolution and split-ergativity: Evidence from Georgian

Dato Abashidze[1], Rusudan Asatiani[2]
[1]The Leibniz-Centre General Linguistics (ZAS), Germany
[2]TSU, Georgia

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Research on pronoun resolution, predominantly conducted on well-studied nominative-accusative languages, has revealed that a range of factors influence the interpretation of pronouns. The most discussed factors are topicality, thematic role [1], discourse status [2] and grammatical role parallelism [3]. The latter aspect is present in nearly all types of texts and communicative settings and as such the interest in the role it plays in pronoun resolution has seen a recent re-emergence [4, 5, 6]. When resolving ambiguous pronouns, studies with monolingual German speakers showed a bias for a subject antecedent, while L2 speakers of German (with L1 Georgian) revealed an overall preference for an object antecedent [6]. Another study on Georgian using the same experimental setup [7], but with antecedent sentences in the nominative-dative construction, also revealed a bias towards the object antecedent. However, it has not yet been explored to what extent Georgian speakers apply grammatical parallelism in a morphosyntactic ergative construction while resolving ambiguous pronouns.

In a visual world paradigm study, eye-gaze data were collected from students (N = 46, M = 19.84. years, SD = 2.01) of Georgian Universities. Participants viewed a depicted scene (showing a crow, a boar, and a rock concert scene) and listened to Georgian sentences (e.g., ყვავმაERG ტახიNOM გასართობად როკკონცერტზე დაპატიჟა. ისგულს იჯერებდა ცეკვით /მას ენერგიას ჰმატებდა სასმელი. Translation: The crowERG invited the wild boarNOM to a rock concert. HePROsbj enjoyed the dancing /HimPROobj gave the drink the energy". The antecedents (with subject-object-verb word order) were presented with the case morphology in the noun phrase determining the grammatical role. The effects of grammatical role parallelism were examined through the subject ის "he" or object მას"him" pronoun, in the following sentence. Eye movements to the subject and object antecedents were analyzed from the pronoun onset until the end of sentences.

The preliminary eye-tracking results revealed that participants started to use the pronoun information immediately after its onset. Within the first 300 milliseconds, they preferentially inspected the object antecedent (i.e., wild boar) in both pronoun conditions, which is in line with the previous findings [6, 7].

However, the inspection of the subject and object antecedent as a function of the pronoun started to diverge at approximately 350 milliseconds after its onset. Participants inspected the subject antecedent (i.e., crow) more frequently in the subject pronoun trials compared to the object pronoun trials and vice versa in the object pronoun condition. A follow up offline test only partly supported the eye-gaze patterns, as the offline choices were associated with the more frequent choice of subject antecedent in subject pronoun trials and vice versa. The results are discussed in relation to ergative construction and ambiguous pronoun processing.

References

[1] Schumacher, et al., 2016.
[2] Kaiser, E., Trueswell, J.,C 2008.
[3] Smyth, R. 1994.
[4] Sauermann, A. Gagarina, N. 2017.
[5] Poirier, J., Walenski, M., Shapiro, L.P. 2012.
[6] Abashidze, et al., 2022.
[7] Under review et al. 2023.

The role of lexical-semantics in incremental adjective interpretation

Stavroula Alexandropoulou[1], Nicole Gotzner[2]
[1]UCL, United Kingdom
[2]Osnabrueck University, Germany

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

The type of comparison standard evoked by gradable adjectives (e.g., warm) has been found to affect the availability of their pragmatic inferences (Gotzner et al., 2018). It further divides gradable adjectives into relative (warm) and absolute adjectives (bent): while for relative adjectives the value on the underlying measurement scale serving as a comparison standard is contextually determined, for absolute adjectives this is a context-invariant value (Kennedy & McNally, 2005). Crucially, it has been argued that scalar implicatures (SIs) of relative adjectives (warm->'warm but not hot') are not derived in all contexts because one needs to be able to resolve the standard value for each of the two scale-mates (warm vs. hot), whereas absolute adjectives are more robust SI triggers.

In the present study, we investigate how lexical-semantic factors such as the type of (comparison) standard affect the incremental SI derivation for different adjective types, by means of two experiments. Experiment 1 is an incremental decision experiment (Qing et al., 2018) and Experiment 2 is an eye-tracking experiment, both using the visual world (VW) paradigm and conducted online on PCIbex. We build on Aparicio et al.'s (2015, 2018) VW eye-tracking studies, demonstrating that the lexical processing of relative adjectives hinges on the visual presence of a contrast object that helps fixing the standard (Sedivy et al.'s (1999) referential contrast effect), whereas the processing of minimum-standard absolute adjectives relies solely on linguistic information. In Experiment 1, using a referential communication task and a visual scene, the referring written instruction (`Click on the picture of the warm water with the purple spoon') was revealed incrementally and for every single instruction segment participants had to click on what they believed the intended referent to be. The instruction containing a gradable adjective (warm) was temporarily ambiguous between two referents in the visual scene. In order to disambiguate between the two, one had to draw the SI associated with the adjective of the instruction ('warm but not hot'). Hence, disambiguation in favor of a referent was reflected in a high(er) proportion of clicks on that referent. We manipulated the adjective type and the visual context (contrast vs. no-contrast object). Experiment 1

revealed that the incremental SI computation was facilitated by the immediate visual context for relative vs. minimum-standard adjectives.

Experiment 2 uses the same materials as Experiment 1, and tests whether lexical-semantic differences of relative and minimum-standard adjectives affect the incremental SI computation in the same direction as in Experiment 1, under the assumption that the proportion of fixations on a referent reflects the degree of belief that this referent is the intended one (Degen et al., 2021). Eye-movements are recorded using the PCIbex software (Zehr & Schwarz, 2018) and the WebGazer.js algorithm (Papoutsaki et al., 2016). Pilot data reveal that participants fixate on the target referent faster when there is visual contextual support in the relative adjective condition than without such support, whereas such a difference is smaller for minimum-standard adjectives.

Our findings suggest that semantics and pragmatics are highly intertwined during incremental adjective interpretation.

Split topicalization in Jordanian Arabic: remnant movement analysis

Eman Al Khalaf
The University of Jordan, Jordan

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

This paper empirically and theoretically investigates the syntax of split topicalization in Jordanian Arabic, a previously undiscovered phenomenon. The results of a large-scale acceptability judgment task (n = 463) reveal that Jordanian Arabic displays split topicalization of various categorial types. Building on the labeling framework proposed by Ott (2015), the remnant movement analysis of Muller (1998) and the idea that there exists an IP-internal focus position below T and above the vP phase (e.g., Belletti, 2004), I propose a unified analysis of all the categorial types of ST in Jordanian Arabic, where ST is reduced to remnant movement of a constituent (NP, VP, AP, etc.), from which a subconstituent has undergone focus movement to an IP-internal focus position. The analysis captures all cases of ST in Jordanian Arabic, including recalcitrant data, such as splitting with ditransitives. It also accounts for the fact that ST in JA is only grammatical with bridge-contour intonation: the intonation is read off the syntactic structure in which functional projections such as TopP and FocP are encoded. If the analysis is on the right track, it provides insights into the structure of the left periphery of the vP phase: it lends support to the view that this area could have a more elaborate structure than previously thought and parallel to that of the left periphery of the CP phase.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Discerning similar speech

Terese Anderson, Grandon Goertz
University of New Mexico, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Introduction

The voices of close family members are often mistaken for one another in social gatherings and especially in phone conversations. We examined the recorded voices of a pair of female family members whose speech is often confused with that of the other person.

This research was designed to determine which voice qualities account for the similarity of voicing between family members. We compared the formants of telephone voice recordings to studio microphone recordings, to determine which phonetic qualities could account for one family member being mistaken for another.

This research discusses innovative methods that were employed to determine which voice qualities are indicators of this voice similarity for phone conversations.

Object of Study

This study investigated analytical comparison of voice recording to studio recording and if an explanation for voice similarity could be identified using computer analysis technology. We examined if other parameters could be used to identify speakers.

Methodology

The two female speakers separately read a 98-word paragraph in English in a professional sound booth. Each speaker produced two simultaneous voice recordings: a studio recording and a voicemail recording on a cell phone. This process produced four recordings.

A series of investigatory computer algorithms was used to compute F0, or the pitch for each of the four recordings. The first computation was the determination of the harmonic ratio, which measures the ratio of energy in the harmonic portion of sound to the total energy of the sound. The harmonic ratio will produce a more accurate pitch calculation. The second evaluation was the computation of subharmonic frequencies which are frequency widths that are characteristic to each sound. The third evaluation was the use of power spectrums to provide a plot of a signal's power within given frequency ranges and is used to find the dominate frequency.

Conclusions

Comparing a speaker's studio recording to their phone recording did not produce explainable differences. Phone voicemail recordings have irregularly shifted pitch frequencies to both lower and higher frequencies compared to the studio speech sample, thus distorting the sound.

The voicemail recordings contain formants which vary widely and irregularly from the studio recording. Examination of the differential in formant values for both speakers did not yield a consistent pattern of shifting from the voicemail to studio recording. This kind of variance in formant values does not point to a confident method for voice analysis comparisons.

Comparisons of studio-produced waveforms and attendant F1-F4 formants between speakers were not reliable indicators of speaker similarity.

Changes in the power spectrum show that formant analysis may be unreliable. Power spectrum comparisons for each speaker show 100-fold reduction in power comparing the spoken voice to the voicemail. Spectrum frequency peak values were near to 500 Hz for the voicemail and 1000 Hz for the microphone speech.

In all four speech samples, both speakers had comparable subharmonic ratios, and word space timing for intonation units. These findings indicate that these two qualities may indicative for voice distinction.

Lexical aid as an accommodation in assessing reading comprehension of English language learners with Specific Reading Disability (SRD)

Georgia Andreou, Panagiota Athanasiadou
University of Thessaly, Volos, Greece

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Specific Reading Disability (SRD) constitutes a specific learning disability which affects decoding, fluency and comprehension of written input. The severity of deficiencies in the aforementioned domains exacerbates in the context of second language (L2) learning, depending on the level of transparency of the L2 and the similarities it presents with the first language (L1). Greek learners face significant difficulties in learning English, since their L1 presents a transparent phonological system, whereas English is a non-transparent language (Abedi et al., 2005), rendering reading skills very challenging. Therefore, the performance of Greek SRD learners is significantly lower in reading comprehension assessment compared to their typical peers, which makes the use of accommodations urgent, so that an accurate picture of their knowledge can be achieved (Sireci, Li & Scarpati, 2005). An integral part of reading comprehension in L2 is the knowledge of vocabulary. SRD learners are exposed to fewer texts; therefore their vocabulary is limited (Vermeer, 2001). One accommodation proposed for this inherent deficiency is the use of lexical aid. Limited studies have been conducted on the issue with ambivalent results. The aim of the present study is to investigate the effectiveness of the accommodation of lexical aid in the reading comprehension assessment of SRD learners. 48 typical and 48 SRD students have completed a placement test with the aim of level homogeneity as well as WISC-III and A-test aiming at the specification of their profile. Two forms of a reading comprehension test were administered to four groups. The two conditions (standard and accommodated) were compared in a design of repeated measures type, which also controls the effect of order, where all students were measured in all conditions. The accommodated format included the provision of English equivalents for the challenging words of the texts, which were specified through an application based on corpora. Results showed that the SRD group presented lower performance compared to the typical group; however, their performance improved significantly under the accommodated format in contrast to typical group's performance. Furthermore, the effect of fatigue did not present stable results regarding the two-form order. This finding demands further investigation so that information can be collected on text length and number of

texts used during assessment. The underperformance of SRD learners, specifically in L2 context, constitutes a prominent problem that remains unexplored. Therefore, this study aims at offering significant information on the scientific area of assessment for learners with specific learning disabilities, moving a step further from differentiated teaching to differentiated assessment.

References

Abedi, J., Courtney, M., Mirocha, J., Leon, S., Goldberg, J. 2005. Language Accommodations for English Language Learners in Large-Scale Assessments: Bilingual Dictionaries and Linguistic Modification. CSE Report 666. National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Sireci, S.G., Scarpati, E., Li, S. 2005. Test accommodations for students with disabilities: An analysis of the interaction hypothesis. Review of Educational Research, 75, 457–490.

Vermeer, A. 2001. Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. Applied Psycholinguistics, 22(2), 217-234.

Language input of Cypriot children: analysis with LENA

Paris Binos[1], Loukia Taxitari[2]

[1]Department of Rehabilitation Sciences, Cyprus University of Technology, Cyprus
[2]Department of Psychology, Neapolis University, Cyprus

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

The Language Environment Analysis (LENA Research Foundation, Boulder, CO, USA) is an innovative "talk pedometer" that fuses a wearable audio recording device with automated vocal analysis. According to nativist theories of language acquisition, language input merely acts as a catalyst for the maturation of linguistic competence.

This groundbreaking study, the first of its kind involving Greek-speaking families, was designed to unveil the language habits of families with typically developing children. Within this unique linguistic community, the study endeavours to discern the exact nature of the input children learn from and how it influences language acquisition. How does the recorded child-directed speech (CDS) evolve over developmental time?

The study included four Cypriot Greek-speaking families with typically developing children aged between 6 and 46 months. Four natural interaction sessions were recorded for each child, totalling 214 hours. The audio data was then automatically processed and categorized by the LENA Pro software into eight distinct classes: conversational turn count (CTC), child vocalisation count (CVC), adult word count (AWC), and an automatic vocalisation assessment (AVA), which includes meaningful audio environment, distant audio environment, TV and electronic sounds, noise, silence, and background sounds.

The children's performance and language input conformed to the standard scores of the LENA system for typically developing children. However, the system has limitations as it classifies these vocals as one CV when a child produces prelinguistic vocalisation in a sequence or a single breath. Human coding is necessary to identify specific vocalisations in word or syllable structure. Recent research corroborates the findings of this study. The analysis included descriptive measures and more advanced statistical tests. The results measured the mean of AVA parameters during different time intervals. There were high correlations ($r > .06$) between the amount of parental language, turn-taking, and child vocalisations, indicating the significant influence of adult speech on children's language production. Additionally, regression models with AWC, CVC, and CTC as the dependent variable and the parameters of AVA as independent variables were highly significant, showing the meaningful audio

environment as the most crucial variable in explaining the variability in all dependent variables. Intriguingly, the use of electronic devices also accounted for the variability in child vocalisations but negatively: the more the usage of electronic devices, the fewer child vocalisations were produced.

Therefore, the LENA software can furnish valid and reliable measurements of a child's communication environment in languages beyond English, such as Greek. It provides critical information about adjusting various aspects of language input to aid language development. Through LENA, parents, teachers, and clinicians can easily track the volume and quality of their child's language input and juxtapose it with their child's language development.

Judgment testing of adverb topicalization in irregular wh-questions

Sophie Boes, Cameron Wacker, Ross Klein
University of Wisconsin-Madison, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Background

In current scholarship, irregular wh-questions (IWQs)—wh-questions in English such as what about—are used as diagnostic tools in discourse studies, in which IWQs are followed by a gerund clause or NP (Huddleston et al. 2002). However, recent data extracted from the British National Corpus (BNC) led Li & Lui (2023) to purport that "topicalizable" adverbs like domain, locative, and temporal adverbs can follow IWQs, but "untopicalizable" adverbs like epistemological, predicational, and discourse-oriented adverbs cannot. The current study tests Li and Lui's (2023) claim through a more rigorous method of examination—a contextualized acceptability judgment task (AJT)—to explore the notion that relying on corpora alone to study syntax is problematic.

Method

Data was collected via a 7-point Likert scale AJT (12 contextualized tokens and 24 contextualized fillers) from 51 native speakers of American English. To examine the interaction between IWQ and adverb type, the experiment utilized a 2x2 factorial design that crossed two factors: question type (irregular wh-question vs. wh-question) and adverb type (untopicalizable adverb vs. topicalizable adverb). Table 1 represents the test design with a sample token set.

Table 1: 2x2 Factorial Design

Context: Tim loved to listen to music in the car, but his little brother, a talented piano player, constantly critiqued his taste and asked annoying questions about aspects of the music. This morning, Tim's brother asked:

	Irregular Wh-Question		Wh-Question	
Topicalizable song's sound?	1:	*What about clearly the song's sound?	2:	What is clearly the song's sound?
Untopicalizable song's sound?	3:	(3) What about tonally the song's sound?	4:	What is tonally the song's sound?

Results

We conducted a linear mixed-effects regression analysis on z-transformed ratings with question and adverb types included as fixed factors and subject and item included as random factors. Results indicated significant effects of question type (Estimate = 0.267, SE = 0.132, $p < 0.05$), as wh-questions were rated higher than irregular wh-questions. However, there was no significant interaction effect of question and adverb types, indicating that participants rated irregular wh-questions lower, regardless of adverb type. This contradicts the claim made by Li & Lui (2023).

Discussion and conclusions

A contextualized AJT failed to find the interaction between question and adverb types purported by Li & Lui's (2023) probing of the BNC. This underscores the potential pitfalls of using production data like corpora to study syntax because "absence of evidence does not mean evidence of absence"; in other words, merely because IWQs with untopicalizable adverbs are not frequently found in the BNC does not mean that they are necessarily ungrammatical. Nonetheless, the interaction between question and adverb types warrants further research: If adverbials can appear as a topic in IWQs—a novel assertion that adds to the existing literature on topics and adverbials—then that fact should be corroborated by judgment data.

References

Huddleston, R., et al. 2002. The Cambridge grammar of the English language. Cambridge: Cambridge University Press.

Li, W., Liu, J. 2023. About 'what about': the semantics and syntax of irregular wh-questions in English. Linguistics 61(1). 159–195.

Vowel perception of American English

Antonis Botinis[1], Grandon Goertz[2], Athina Kontostavlaki[3], Terese Anderson[4]
[1]Department of Research and Development, ExLing Society, Greece
[2]Department of Educational Linguistics, University of New Mexico, US
[3]Laboratory of Phonetics and Computational Linguistics, National and Kapodistrian University of Athens, Greece
[4]Radiation Safety, University of New Mexico, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Object of the study

In this study we compared the perception of New Mexico American English (NMAE) vowels by native American and native Greek speakers. We examined if all eight monophthongal vowels were descriminated equally by both English and Greek speakers. We also inquired if Greeks discriminate better the NMAE vowels which are closer to the acoustic structure of the Greek five vowels.

Description and theoretical context

The pronunciation of vowels in the English-speaking world varies considerably. In addition to country and related dialectal variability, there is no consensus about the number of vowels in American English or the respective transcription. Furthermore, the IPA (International Phonetic Alphabet) is not the standard phonetic transcription and different authors and dictionaries use a variety of different phonetic-related symbols.

Methodology

Speech data was acquired from one male and one female New Mexico American (NMA) native speaker with key words containing the monophthong vowels/iː, ɪ, e, æ, uː, ʌ, ɔ, ɑː/ in the carrier phrase "He said _____ slowly". Each recorded carrier phrase from each speaker was combined with either itself or one of the other seven carrier phrases, producing eight stimuli pairs for each native speaker, making possible 16 different speech pairs. Two forced-choice discrimination experiments were then carried out. In the first experiment, native New Mexico American listeners were asked to choose between "same" or "different" among pairs of same or different New Mexico American English monophthong vowels. In the second experiment, likewise, native Greek listeners were asked to carry out the same task.

Twenty four English and nine Greek listeners participated in the study.

Results and conclusions

The New Mexico American listeners discriminated correctly over 88% all vowels except for the vowels /ɔ/~/ɑ/, as in /bɔt/ "bought" and /bɑ:t/ "bot" pairs. This means that the distinction between these two vowels has been neutralised in this particular dialect of New Mexico English. The discrimination rates regarding male and female listeners for both male and female speakers do not show any noticeable differences.

The Greek listeners discriminated the vowels above 61% except for the pair /bɔt/~/bɑ:t/, i.e. in a similar way to the New Mexico American listeners. The results for the Greek listeners also show a different pattern in that the American English vowels were discriminated at either a higher discrimination rate or a lower rate. The discrimination rates indicate that the Greek speakers tended to discriminate vowels in one of three phonological groups.

Our summarized results are shown in these two tables

:

Table 1. Correct discrimination rates of 4 male listeners and 20 female listeners, a total of 24 listeners, for American English as a function of one male and one female speaker.

	beat	bit	bet	bat	boot	bought	bot	butt
beat	91%							
bit	93%	89%						
bet	90%	90%	91%					
bat	96%	94%	90%	90%				
boot	94%	99%	98%	95%	91%			
bought	99%	93%	96%	89%	94%	90%		
bot	99%	94%	98%	98%	94%	11%	88%	
butt	96%	95%	93%	91%	95%	99%	94%	91%

Table 2 shows the Greek data, which consists of two speakers, one male and one female, and 9 listeners, 2 male and 7 female ones. The results indicate that the discrimination rate forms three? groups.

	beat	bit	bet	bat	boot	bought	bot	butt
beat	80%							
bit	81%	61%						
bet	100%	97%	83%					
bat	94%	97%	94%	77%				
boot	100%	97%	94%	94%	86%			
bought	91%	100%	88%	66%	97%	73%		
bot	94%	97%	88%	66%	97%	31%	67%	
butt	100%	97%	100%	75%	100%	84%	66%	91%

Linguistic vs. tactile experience in object classification

Letizia Cerqueglini
Tel Aviv University, Israel

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

I disambiguate the role of language vs. tactile experience in object conceptualization through linguistic and cognitive testing of frames of reference (FoRs) in 12 sighted (TNA)/12 blind (BTNA) Traditional Negev Arabic speakers. FoRs are coordinate-systems projected onto spatial arrays to locate any object (Figure, F) in relation to another object (Ground, G). FoRs are object-centered, based on inherent facets of G; ego-centered, based on the coordinates projected by the speaker; geocentric, based on external coordinates (cardinal directions). TNA is a cluster of North-West Bedouin Ḥijāzi Arabic dialects spoken in the Negev (Israel) by women/men over age 75 without formal education. BTNA speakers received no special instruction/training of the type modern societies provide special needs individuals. Following Cerqueglini (2023), TNA speakers select among linguistic FoRs based on values culturally attributed to Gs while cognition is only geocentric. BTNA linguistic and cognitive FoRs align with TNA. Nonetheless, BTNA speakers prefer ego-centered over object-centered and geocentric FoRs with highly frequent (knife/cup/glass) and not easily graspable Gs (glass/dish), using ego-centered right/left representations, which are absent from TNA, where 'right'/'left' only distinguish between hands. BTNA cognitive data could be interpreted as language-driven or represent the effects of non-visual, sensory (tactile, motor, and so forth) experience. I tested the effects of the tactile channel vs. linguistic constraints on linguistic and cognitive object conceptualization. Novel objects without nouns neutralize linguistic effects. The tactile channel was the only experiential channel possible/permitted. TNA speakers were blindfolded. FG stimulus arrays were arranged one after the other on a table accessible to the informant's touch, attached to a base so as not to be moved. F was a ping-pong ball. Gs were 12 novel objects with different values of geometric, functional, and cultural parameters (with/without pseudo-handles, pseudo-body parts) that resembled real objects in shape and material. Informants were individually tested on the same arrays twice—once using the dominant hand (DH) and once the non-dominant hand (NDH). Maximum reaction time was five seconds. Linguistic test: for each array, informants were asked in their language to say where the ball was in relation to the other item. Cognitive test: Recall Memory Task (Levinson 2003: 154–160).

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

In 23% of TNA and 14% of BTNA linguistic responses, informants attributed an existing noun to G (62% of TNA and 38% of BTNA DH-responses, 35% of TNA and 24% of BTNA NDH-responses). With named Gs, linguistic FoR was G-based (TNA 83% and BTNA 21% DH-responses; TNA 17% and BTNA 7% NDH-responses). Geocentric FoR was applied in all other cases. BTNA diverged from TNA G-based FoR selection, as the ego-centered FoR took over the others, frequently in DH-responses with asymmetric, easily graspable Gs (69% vs. 16% of NDH-responses). TNA cognition was always geocentric. BTNA cognition was exclusively geocentric in NDH-responses and ego-centered in 57% of DH-responses, within which 67% were not elicited from Gs named in the linguistic test, demonstrating that tactile experience drives the expansion of ego-centered FoR in both BTNA language and cognition.

Role noun processing and gender stereotypes in Greek

Despoina Chalyvidou, Andrea Weber
Eberhard Karl University, Tübingen, Tübingen, Germany

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Previous studies have shown that information about the stereotypical gender associated with role nouns is incorporated into the mental representation of these nouns and is activated as soon as a role noun is encountered (e.g., Oakhill et al., 2005). When hearing the word "nurse" or "electrician", information about the person being putatively female or male will be activated. Consequently, stereotypically incongruent information (female electrician or male nurse) will slow down processing (Siyanova-Chanturia et al., 2015). In languages with grammatical gender, grammatical cues can also mark the gender of a character introduced by a role noun as male or female. While most Italian nouns are marked for gender, studies examining role noun processing in Italian have mainly employed bigender nouns with no gender marking, thus only the effect of stereotypes has been investigated for these nouns (Cacciari & Padovani, 2007; Siyanova-Chanturia et al., 2015). The present study aims to extend previous findings by examining role noun processing for the first time in Modern Greek using nouns of common gender that are morphologically marked for masculine gender, to investigate the interplay of grammatical and stereotypical information during processing.

Our experimental paradigm is based on the study by Cacciari and Padovani (2007): native Greek adults are primed with occupational nouns of common gender with a masculine ending and a male or female bias (e.g., iðravlikós "plumber", esθitikós "beautician"), followed by a masculine or feminine pronoun (aftós "he" / aftí "she"). The prime-target pairs are stereotypically congruent (iðravlikós – aftós / esθitikós - aftí) or incongruent (iðravlikós - aftí, esθitikós - aftós). Forty role nouns were selected as primes: 20 stereotypically female and 20 stereotypically male. The participants' task is to decide the gender of the following pronoun, and reaction times and accuracy are being measured. The results will indicate if the findings reported by Cacciari and Padovani (2007) can be replicated when the stimuli comprise nouns with gender marking, albeit not being informative about referential gender. If stereotypes guide processing, then our results will be in line with those of Cacciari and Padovani, with slower reaction times for the incongruent stimul. If, however, morphosyntactic cues further impact processing, then the masculine morphology of the nouns will either lead to male-biased processing

despite the stereotypical bias conveyed by the nouns, or, at least, it will modulate the strength of the stereotypes. Our research will provide insight into the interaction of grammatical gender and gender stereotypicality during processing for Modern Green, which might even have implications for the use of these words, requiring the adoption of strategies towards a more gender-fair and inclusive language.

References

Cacciari, C., Padovani, R. 2007. Further evidence of gender stereotype priming in language: Semantic facilitation and inhibition in Italian role nouns. Applied Psycholinguistics, 28(2), 277–293. https://doi.org/10.1017/S0142716407070142

Oakhill, J., Garnham, A., Reynolds, D. 2005. Immediate activation of stereotypical gender information. Memory & Cognition, 33(6), 972–983. https://doi.org/10.3758/BF03193206

Siyanova-Chanturia, A., Warren, P., Pesciarelli, F., Cacciari, C. 2015. Gender stereotypes across the ages: On-line processing in school-age children, young and older adults. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.01388

Acceptability of agreement variants of there existential construction

Vatcharit Chantajinda
University of Wisconsin-Madison, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Previous corpus-based research suggests that "there's" followed by a plural noun, such as "there's ants in the syrup," is highly frequent, especially in spoken English (e.g., Riordan, 2007). This occurs despite the disagreement between the cliticized copula and the grammatical subject "there" which is linked to the postverbal subject NP sharing number agreement through chain (Haegeman & Guéron, 1999). Evidence from experimental sociolinguistics (Hilton, 2016) also found that "there's" is not perceptually different from the standard form "there are".

This study examines the acceptability of "there" existential construction and its variants to explore syntactic variation by using an experimental linguistic method. The aim of this study is to investigate if data obtained from a judgment task aligns with previous studies that used production and perception data. Twenty native speakers of English completed a seven-point scaled acceptability judgment task administered with a 2x2x2 factorial design, including three two-level factors: copula realization (full vs. cliticized), agreement (agree vs. non-agree), and quantifier (bare NP vs. a lot of NP). The postverbal subject NP is plural in all test items. Sixty-four items were presented on Qualtrics, pseudorandomized, and counterbalanced for the number of acceptable and unacceptable conditions, with 16 experimental items and 48 fillers. Results from z-scores and a linear mixed-effects model in R showed that, considering "there's", "a lot of NP" received higher rating scores than "bare NP", aligning with previous corpus-based research that found a significant difference between the two conditions (Riordan, 2007). However, the present study found no effects of the quantifier. Excluding quantifier, copula realization and agreement were analyzed using the same model. An interaction between copula realization and agreement was found. Paired comparisons indicated that the four forms all significantly differed from each other with "there are" (full, agree) rated highest, followed by "there're" (cliticized, agree), "there's" (cliticized, non-agree), and "there is" (full, non-agree), respectively.

The findings thus reveal that "agree" results in higher acceptability as the two forms were rated more acceptable than the two "non-agree" forms. Copula realization, however, affected the two agreement conditions differently. In the "agree" forms, the full copula form ("there are") is more acceptable than the

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

cliticized one ("there're"). In the "non-agree" form, the cliticized form ("there's") is more acceptable than the full copula one ("there is"). The findings suggest that, despite its frequency found in literature, "there's" is still less acceptable than the "agree" forms, indicating a strong connection between "there" and the post-verbal NP. However, task effects should be considered as this study used a written AJT while "there's" is frequent in spoken language.

References

Haegeman, L., Guéron, J. 1999. English grammar: A generative perspective. Blackwell.

Hilton, K. 2016. Nonstandard agreement in Standard English: The social perception of agreement variation under existential there. University of Pennsylvania Working Papers in Linguistics, 22(2), 61-70.

Riordan, B. 2007. There's two ways to say it: Modeling nonprestige there's. Corpus Linguistics and Linguistic Theory, 3(2), 233-279.

Processing speech acts: spoken communication for aircraft maintenance

Savvas Chatzipanayiotidis, Christina Alexandris
National and Kapodistrian University of Athens, Greece

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

We focus on the data obtained from Greek spoken technical texts for aircraft maintenance and technical support. The data is presented corresponds to the analysis stage for subsequent processing and integration in Human-Computer Interaction applications involving speech.

The identification of speech act types and speech act content constitutes the basis for the construction and implementation of a flexible template-based, slot-filling framework for multilingual spoken technical texts in the domain of aircraft maintenance for its direct deployment or its subsequent integration in neural network approaches and/or as training data. In particular, the presented slot-filling framework is targeted to process spoken utterances from the domain of aircraft maintenance and technical support - without relying to a considerable extent on the quality of the Speech Recognition system (ASR).

Speech act recognition, in combination with the recognition of the appropriate word groups contributes both to the speed and to the correctness of processing spoken input. In addition, the present approach also focuses in the minimum possible size of the sets in keyword recognition, taking into account the varying degrees of quality of ASR systems. The features and constitute basic elements of the proposed template-based, slot-filling framework which is based on keyword recognition with keywords in the form of tuples (x,y). In particular, the present approach is based on the mandatory recognition of two keywords in complimentary relation (x,y). According to the evaluation performed on the data presented, this relation is observed to ensure the correct identification of speech act type as well as the correct identification of content type in the spoken utterances. Furthermore, the recognition of two keywords in complimentary relation accounts for cases in which speech act type and content type coincide or overlap, a possibility in some types of commands and utterances concerning safety regulations and alerts.

The keywords extracted from the content of each spoken utterance are integrated to the "generic-intent" slot of the template-based, slot-filling framework. The information content of the "generic-intent" slot is linked to the respective Speech Act. The keywords are related to 620 expressions from approximately 2800 spoken utterances extracted from data was collected both from and available datasets of transcribed Greek spoken technical texts and

from transcribed spoken dialogs for aircraft maintenance and technical support (Greek). The 620 expressions are mapped into 21 distinct tasks and related Speech Acts presented. The Speech Acts related to the 21 distinct tasks are divided into three general categories, namely "Command-Question", "Information-Answer" and "Emergency" and contain distinct speech act subgroups (21), listed in order of frequency and demonstrated in the present research. Most of the respective (620) keywords constituting the actual information type belong to the following general groups: "Wh-Questions", "Quantitative Information", "Other Questions" and "Written Documentation".

The mandatory recognition of two keywords in complimentary relation (x,y) observed to produce correct results for Greek spoken Technical texts is intended to be compared to similar data, if available, in English and German.

Prosodic description of statements and questions in Moore

Laetitia Compaoré
STL-Savoirs, Textes, Langage UMR 8163, France

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

This study examines the use of prosody in the identification of statements and yes/no questions intonation in Moore (a tone language mainly spoken in Burkina Faso).

The analysis describes the phonetic realization of intonation in the two types of sentences. The main issue is to find out the acoustic cues which permit the identification of each type of sentence.

3 native speakers of Moore have each of them recorded 8 statements and 8 yes/no questions. So the corpus of analysis is made up of 48 sentences. All of the sentences contained one prosodic phrase i.e. a sentence with 5 to 7 syllables.

The prosodic description was focused on the acoustic indices located at phrases boundaries, in other words at the end of the sentences.

The corpus was also resynthesized in order to test that the acoustic parameters identified in the analysis were actually perceived. 30 speakers of Moore participated to the perception task.

The results confirm, as in Rialland (2004) and Kinda (1984), that yes/no question have a falling intonation contour in Moore.

They also suggest that yes/no questions have wider register range and longer final syllables duration than statements in Moore.

References

Anyanwu, R.-J. 2008. Fundamentals of Phonetics, Phonology and Tonology. Peter Lang.

Creissels, D. 2003. Some remarks on tonal processes at utterance boundaries. Frankfurter Afrikanistische Blätter, vol. 15, pp. 49-58.

Goldsmith, J. 1976. Auto segmental phonology, Ph.D. dissertation MIT.

Kaboré, R. 1985. Essai d'analyse de la langue mooré (Parler de Wàogdgò : Ouagadougou). Thèse, Université Paris 7.

Kinda, J. 1984. Dynamique des tons et intonation en mooré, Thèse. Paris: Université de la Sorbonne-

Martin, Ph. 2009. Intonation du français. Paris: Armand Colin.

Morel, M.-A., Danon-Boileau, L. 1998. Grammaire de l'intonation, l'exemple du français oral. Paris:

Rialland, A. 2004. Tonologie africaine et modélisation prosodique", Théories linguistiques et langues sub-sahariennes, L'Harmattan, Paris , prefinal version

Prosody and pragmatics of vocative in spoken Italian

Emanuela Cresti, Massimo Moneglia
University of Florence, Italy

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Vocatives have a reduced lexicon: proper names, nouns of family and professional roles (mom, dad, professor, doctor, boss, man), and appellative adjectives (dear, honey, rogue, her majesty, his holiness). In Italian (Mazzoleni, 1995), includes nouns preceded by the third-person demonstrative (quel giovane) or followed by the first possessive pronoun (figlio mio, amica mia).

Grammars give little space to Vocatives. It is supposed that the semantics of vocatives identify as deictic devices the Addressee by referring to him. Vocatives can play a broad functional spectrum. Quirk et al. (1985) distinguish between "a call drawing the attention of the person addressed, singling them out of others in hearing" and "an address expressing the speaker's relationship or attitude to the person addressed". Biber et al. (1999) also consider the maintenance and reinforcement of social relationships and focus on the initial or final distribution in the utterance. The final most frequent position "stresses the social relationship", while the initial is dedicated to "prompt attention". According to Mazzoleni (1995), the prompting function is used in isolation, waiting for a reaction from the addressee. Beyond these observations, little attention is given to prosody.

Based on the Italian C-ORAL-ROM (Cresti & Moneglia, 2005) stored in the Database IPIC of the information structure (Panunzi & Gregori, 2012; Panunzi &Mittmann, 2014), the paper will address three questions: a) what the actual range of vocative interpretations in spoken Italian is; b) how these interpretations can be decided; c) what does constraint their strong lexical selection.

We will argue that prosodic performance specifies the function of vocatives, which can be framed within the sharp distinction between Illocutionary functions and dialogue regulation functions foreseen within Language into Act Theory (Cresti, 2000; Moneglia & Raso, 2014; Cresti & Moneglia, 2018).

When vocatives perform an autonomous speech activity of Calling (illocution), they give rise to a simple utterance. Its pragmatic function is to open the communication channel, and the prosody varies according to the position of the addressee (proximal vs. distal) (Cresti, 2020).

Calling illocution can also serve to activate the addressee in connection to a specific speech activity (functional recall). In this case, the channel should be

open, and the calling act is patterned by prosody together with a second directive illocutionary unit (Illocutionary Pattern) (Panunzi/Saccone, 2019)

When Vocative expressions are not autonomous, they constitute an optional Information unit within the utterance, working as a communicative support. The role is to stress the social cohesion with the addressee and only secondarily to get his attention. By preferences, these units, called Allocutive in L-AcT, conclude the utterance. Allocutives should be distinguished from other dialogic units (Conatives), specifically dedicated to getting the addressee's attention, for their defocused prosody and lexical constraints.

Corpus observation supported by competence judgments can explain that in vocative, proper names lose the referential value of rigid designators (Kripke, 1980) in favor of a designatory reading required by their function. For this reason, contrary to what might be expected, deictics and referential definite descriptions can never fill calling utterances or Allocutive units.

Bengali ESL learners' acquisition of complex prepositions

Shreya Datta, N.P. Sudharshana
Indian Institute of Technology, India

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

This study intends to examine how ESL (English as a Second Language) learners of different proficiency levels, having Bengali as their L1, acquire the path of traversal- Through and the polysemous preposition Over. These prepositions are known to create learnability issues in ESL/EFL (English as a Foreign Language) contexts: The former preposition, referred to as 'Path of Traversal,' has complex geometric and functional specifications regarding the ground and the path traversed, while the latter is highly polysemous. In addition, there are crosslinguistic differences between Bangla and English in encoding comparable spatial scenes and the lexicalization options available. According to these differences in the lexical packaging of path information, Talmy (1985) classified languages into two types: satellite-framed languages (S-languages) and verb-framed languages (V-languages). English falls under the S-languages, where the manner and motion are encoded in the verb, while the path information appears outside the verb in a satellite, i.e., a prepositional phrase. (E.g. the bottle floated out of the cave). Bengali, on the other hand, is a part of the V-framed languages, where the verb encodes the motion and path, while the manner information is expressed in gerunds or preposition phrases or completely omitted. (E.g. botol-ta guha theke bheshe berolo 'the bottle left the cave floating') (Cadierno & Lund, 2004, p. 142–143).

In this cross-sectional study, our aim was to first determine the proficiency level of the participants through the Cambridge English Language Assessment test. This was followed by two comprehension tests and a production test. In the comprehension test, one was a match the sentence with the appropriate picture task. Here, a sentence along with two pictures was presented to the participants. The picture that one thought determined the meaning of the sentence was to be selected. The second comprehension test was a grammaticality judgment task with a 5-point Likert scale. The participants had to judge the degree of grammaticality of the sentences by selecting the numbers from 1 to 5. Here, 5 represented the most grammatical, and 1 represented the least grammatical. The production test was a picture story that was adapted from the test material "Monkey Book" by Stringer (2005). This story had instances of the target prepositions that the characters in the story encounter on their way.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

There are two issues of concern in the context of second language acquisition that we want to address through this study. The first is the learner cognitive system, which might pose challenges in acquiring semantically complex prepositions, especially at lower levels of proficiency. The second is the crosslinguistic differences, where the lexicalization of the concept of these prepositions might lead to learnability issues.

Person agreement with coordinated subjects in Russian

Tatiana Davidyuk
Lomonosov Moscow State University, Russian Federation
Institute of Linguistics RAS, Russian Federation

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

In Russian, in the case of agreement with coordinated subjects the resolution rules and person hierarchy apply. However, the person hierarchy in Russian can be violated in certain cases. An obvious case of violation is represented by (O)VS word order constructions, where agreement can only occur with the nearest conjunct. Additionally, in the Russian National Corpus we can find several instances of 3rd person plural agreement for coordinated subjects, which include the personal pronoun ja 'I' or ty 'you'. Previously, the possibility of such agreement was experimentally demonstrated for German and Dutch languages (Timmermans et al. 2004). Thus, personal agreement with coordinated subjects in Russian demonstrates variability.

The main focus of my experimental study is the word order. My study consists of four experiments. In the first pair of experiments, coordinated subjects of the form ja/ty i X 'I/you and X' (X is a masculine proper name) were presented. The order of the subject and predicate was varied in the sentences for these experiments. In the second pair of experiments, I fixed the word order by using the VS word order, but I changed the order of the conjuncts. All sentences used the non-past tense, which in Russian requires person and number agreement. All experiments included fillers, with a 1:1 ratio to the experimental sentences. The fillers were divided into grammatical and ungrammatical ones, where the ungrammatical fillers contained errors in agreement or prepositional government. The methodology of the conducted experiments involved assessing the acceptability of sentences using a 7-point Likert scale. The participants were required to evaluate four agreement patterns: 1st / 2nd person singular, 1st / 2nd person plural, 3rd person singular and 3rd person plural. For statistical analysis, I employed regression analysis using a linear mixed model and Tukey's method for multiple pairwise comparisons.

In all experiments, the highest ratings were given to agreement according to the resolution rules – 1st/2nd person plural agreement. However, the word order was found to be significant for its ratings. With the VS word order, agreement with the nearest conjunct becomes possible, but the order of the conjuncts is important. An unexpected result is the ratings for 3rd person plural agreement, which were significantly higher compared to ratings for

ungrammatical fillers and ratings for agreement with the last conjunct. It is noteworthy that third-person plural agreement arises even with coordinated subjects where both conjuncts are 1st and 2nd person pronouns (Belova 2022).

Additionally, we conducted similar experiments, but with sentences in the past tense, which in Russian do not agree in person but agree in gender and number. The detailed results of all our experiments and their contribution to theoretical syntax will be discussed in the presentation.

References

Belova D. Evaluation and reading time of predicate agreement with conjuncts. ExLing 2022 Paris: Proceedings of 13th International Conference of Experimental Linguistics.

Timmermans M., Schriefers H., Dijkstra T., Haverkort M. 2004. Disagreement on agreement: person agreement between coordinated subjects and verbs in Dutch and German. Linguistics 42(5), 905-929.

How individuals with Down syndrome understand ambiguous quantifiers

Sarah Dolscheid[1], Isabel Neitzel[2]
[1]University of Cologne, Cologne, Germany
[2]TU Dortmund University, Dortmund, Germany

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Background
Individuals with Down syndrome (IDS) display difficulties in understanding quantifiers but also in dealing with small numerosities. To better understand what may underlie these difficulties, we examined how IDS interpret the German quantifier eine ('a/one'). This expression is ambiguous as it can refer to the exact numerosity of 'one' but also serves as the indefinite determiner 'a', hence posing a particular challenge for language acquisition.

Methods
36 German-speaking IDS (mean chronological age: 13;09 years, mean mental age: 4;10 years, 19 female) were tested in a truth-value-judgment task (TVJT). Stimuli consisted of a white plastic bowl and small plastic strawberries. IDS were either presented with no strawberry, exactly one strawberry, or two strawberries while being asked 'Is there eine ('a/one') strawberry in the bowl?' We additionally compared performance of IDS to that of typically developing children and adult speakers of German as reported in Dolscheid and colleagues (2019).

Results
Like typically developing children and adult speakers of German, 100% of the IDS correctly rejected the claim that there was eine (a/one) strawberry in the bowl when in fact there was none. Furthermore, 100% of the IDS correctly accepted that there was eine strawberry in the bowl when presented with exactly one token of a strawberry. When asked for eine (a/one) in the context of two strawberries, 83% of the IDS accepted this claim, thus showing no upper bounded interpretation of eine. Conversely, 17% rejected two strawberries as a correct instance of eine. When directly comparing IDS's performance to that of typically developing children and adult speakers of German, a logistic regression revealed that IDS significantly differed from adults in terms of their upper bounded interpretation of eine (z.ratio = -3.924, $p < 0.001$). Unlike adults who predominantly rejected two strawberries as a correct instance of eine (68% rejections), the same did not apply to IDS. By

contrast, IDS's interpretation of eine was comparable to that of typically developing children (11% rejections), and there was no significant difference between the two groups in a logistic regression (z.ratio = -0.72, p = 0.72).

Discussion and conclusion

We found that IDS understand the quantifier eine in a way that is comparable to that of typically developing children. That is, IDS interpret eine as lower-bounded (not zero) and also display knowledge that eine refers to the quantity of one. Critically, however, just like typically developing children, IDS do not yet show an upper bounded interpretation of the term eine. That is, IDS often accepted two tokens as a correct instance of eine since this term is ambiguous in German. Taken together, our findings demonstrate that ambiguities in language can contribute to difficulties in how IDS come to interpret numbers.

References

Dolscheid, S., Schleussinger, F., Penke, M. 2019. When "one" can be "two": Cross-linguistic differences affect children's interpretation of the numeral one. Journal of Numerical Cognition, 5(3), 301–313. https://doi.org/10.5964/jnc.v5i3.199.

Links between children's visual attention and language production

Sarah Dolscheid, Barbara Zeyer, Martina Penke
University of Cologne, Germany

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Background

The way speakers produce an utterance is closely linked to their allocation of attention. For instance, when asked to describe a picture with two characters, speakers are more prone to start an utterance with a visually cued character which is in the spotlight of attention (e.g., Gleitman et al., 2007). However, while such links between attentional orienting and language production have been attested for adult speakers, this relationship has not yet been investigated in children. In the present study, we sought to address this open issue by testing 4- to 5-year-old children in a picture description task while their gaze patterns were monitored via eye-tracking.

Methods

Twenty-seven German-speaking children (13 female, 14 male) were asked to describe pictures of two characters presented next to one another. Children were asked to describe the characters in terms of a conjoined noun phrase (e.g., a fisher and a farmer). We manipulated children's allocation of attention by means of a brief visual cue presented in the place where the left character was about to appear. The cue consisted of a small, red circle presented for the duration of 700 milliseconds. In a baseline condition, no cue was presented.

Results

A generalized linear mixed effects logistic regression model revealed that visual cueing was highly effective in modulating children's attention, as reflected by a significant increase in first fixations to the cued character compared to a baseline without cueing, z.ratio = 6.85, $p < .0001$. At the same time, children were also more likely to start their utterances with the cued character compared to baseline, z.ratio = 2.91, $p = .004$. Children's first fixations to a character also correlated significantly with their propensity to first mention that character, demonstrating that children are more inclined to produce an utterance with an entity that is in their focus of attention, $r = .51$, $p < .01$. Additionally, we examined whether children's first looks were predictive of their order of

mention. We found that children's order of mention depended on where they looked first, as revealed by a significant influence of first fixation (left vs. right character) on children's propensity to first mention a character, $\chi 2 = 54.73$, $p < .001$.

Discussion and conclusion

We found that children's first looks to a character were predictive of their order of mention, showing that children were more likely to first mention a character when they had also initially looked at this character. Furthermore, by experimentally manipulating children's visual attention, we found that children were more inclined to start an utterance with an entity that was in their spotlight of attention. Taken together, our findings provide first evidence of close links between visual attention and language production in children.

References

Gleitman, L.R., January, D., Nappa, R., Trueswell, J. C. 2007. On the give and take between event apprehension and utterance formulation. Journal of Memory and Language, 57(4), 544-569.

Examining acceptability judgments in human and AI utterances

Macy Floyd
University of Wisconsin - Madison, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Experimental syntax and semantics have a long tradition of using acceptability judgments to isolate and describe features of speech (Sprouse & Almedia, 2011; Sprouse, 2018; inter alia). These native speaker judgments have been especially interesting when used to compare and evaluate structures of different languages (Linzen & Oseki, 2018; Sprouse et al., 2016; Wierzba et al., 2023). As Artificial Intelligence (AI) has gained popularity, and with recent questions about the inclusion of AI in academic settings, it is important to look at acceptability of assorted syntactic traits as produced by a human speaker versus AI (Chung et al., 2023). Thus, the goal of this study was to assess college-aged students' acceptability judgments when hearing different syntactic structures as produced by a human versus AI.

This quantitative research adapted sentences from Sprouse et al. (2016) and asked participants to use a 7-point Likert scale to rate acceptability of human and AI utterances. Participants were randomly assigned one of four pseudo-randomized, counterbalanced surveys. As is standard practice in experimental syntax, each survey had 32 test items with 16 that were utterances produced by a human, and 16 produced by AI. Participants responded in Qualtrics, then results were analyzed using R. Two-sided T-tests were run to identify statistical significance with a p-value of 5%. In total, 98 responses were collected from college-aged students who were currently enrolled in a university. To avoid confounding variables, the human and AI voices were perceived as both the same age and gender.

Overall, participants who submitted responses seemed to greatly prefer human utterances of the target items over utterances produced by AI. P-values of 2.82% and 2.30% were found for comparisons of all participants and native English-speaking participants as respective groups, which indicates a considerable preference for utterances produced by a human voice over those produced by AI. Statistical significance was similarly demonstrated for comparisons that were constrained to the college-aged group (18-24-year-olds), suggesting that the results were likely not influenced by participant age. Lastly, the statistical significance threshold was not reached when participants were restricted to currently enrolled university students, a potential area for future research.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

The goal of this research was to compare target item production using acceptability judgments, and the overall analysis of this data seems to clearly suggest a preference for human production over AI within this subset of the population. While the results are potentially noteworthy when evaluated alone, this strong preference for human speakers over AI may have effects on other fields of linguistic concern including speakers' acceptability judgments of mistakes, errors and production, the use of AI in L2 classroom instruction in language settings, and using AI in both academia and customer service. This study provides an initial look at using AI in acceptability judgments within experimental syntax and is important to the field as experimental syntax has traditionally used similarities and differences among various formations as the foundation for creating working theories.

Eye movements and reading comprehension in the Greek language while listening to music: an eye tracking study

Maria Gkantaki, Georgia Andreou
University of Thessaly, Greece

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

In the present study, students from the University of Thessaly, Greece, where this research is being conducted, read four different texts in the Greek language under four conditions: while listening to music they prefer, while listening to music they do not prefer, while listening to a recording of noise from a cafe, and finally in silence. After reading each text, they answered a reading-comprehension test for each one. The eye movements of the participants were recorder in all four conditions in order to understand the intersections between reading comprehension in the Greek language and visuality. Also, they were asked to answer a questionnaire about their music and study habits, and a questionnaire in which they assessed their reading attitudes during the experiment. Since previous studies have shown very different results regarding reading comprehension in the presence of music and noise, it is difficult to predict how preferred and non-preferred music as well as noise from a cafe affect reading comprehension in the Greek language. Nonetheless, we propose the following hypotheses: (1) Reading comprehension is enhanced by preferred music; (2) Reading comprehension is negatively affected by non-preferred music and noise from a cafe.

A main effect for reading-comprehension scores reveals that participants score significantly lower after having listened to their non-preferred music while reading, compared to reading in silence. No significant effects are found between the other conditions. This result is a consequence of the participants not being aware that their reading processes are disrupted by a non-preferred musical background. They do not make the necessary changes to the processes involved in reading that are required to compensate for the increased cognitive load caused by reading a text while listening to music they do not prefer and which distracts them. The results are discussed in terms of visual patterns, reading performance, and reading attitudes.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Graded plurality in morpheme representation influences agreement processing

Kalle Glauch
Ruhr-Universität Bochum, Germany

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Recent advances in research on dependency formation suggests that linguistic feature representations in memory are graded, leading to variations in base-line activation based on feature prominence, with more prominent features exhibiting stronger activation (Engelmann et al. 2019).

I propose that the nominal plural morpheme in German, realized through various allomorphic variations differing in prominence, is such a graded feature. I define the prominence of a plural-suffix as its distinguishability, specifically, the number of additional phones it comprises. I posit that plural-suffixes with more phones enhance perception of plurality by amplifying the distinction between singular and plural forms, thereby increasing the morphemes base-line activation.

In regard to agreement-processing, the marking and morphing framework (Bock et al. 2001) predicts the increase in base-line activation of the plural-morpheme to enhance the competition between singular and plural in the continuous number valuation of the whole subject-DP, leading to different processing patterns.

The current study examines the processing patterns focusing on plural subject-verb-number-agreement with collective-noun-constructions manipulated by the plural-suffixes -en (+2 phones), -e (+1 phone), -s (+1 phone) and -ø (+0 phones) of the local noun (the noun in the genitive modifier of the collective construction) in German.

The study uses a RSVP speeded-grammaticality judgement procedure, sequentially displaying each word for 425ms. The study follows a 1x4 repeated measure design, manipulating the plural-suffixes of the local noun. Each item is a complex sentence with a collective-construction subject and a sentence-final plural verb. To prevent lexical confounds, pseudo-nouns controlled for letter and syllable count are used as local nouns.

1)	(Peter	weiß,	Dass	Eine	Vielzahl	Der (the)	Schleimanten	Bier trinken.
	‘Peter	knows	that	a	multitude		Schleimanten	beer drink.’

The participant's task is to judge as quickly and accurately as possible whether the sentence is grammatical after the last word is presented. Both the judgements and the reaction times are measured as response variables.

For the given task, the increase in competition in the subject`s number valuation from more prominent plural-suffixes is predicted to be reflected in a higher acceptance of plural-agreement and higher reaction times for judgements due to additional cognitive effort for number feature selection during dependency formation.

70 participants, prescreened for German as native language, were recruited via Prolific. Each participant provided responses to four items per condition.

The reaction time data was analyzed using a LMM, predicting the reaction time by plural-suffix. A pairwise comparison of conditions revealed significant differences between the overt plural-suffixes -en, -e, -s and the covert plural-suffix -ø. However, no significant difference was observed between the overt plural suffixes with 2 phones and 1 phone.

The Grammaticality-Judgement data was analyzed using a GLMM, predicting judgements by plural-suffix. The same pattern of significant differences as for the reaction time data was found.

The results suggest a binary difference in the influence of number morpheme representation on agreement-processing, with a distinction between overt- and covert plural-suffixes. The data is compatible with representational approaches of number-agreement-processing like marking and morphing but poses a challenge for retrieval-based accounts.

Vowel space and centroid values for three formants

Grandon Goertz
University of New Mexico, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Object of the study
Vowel spaces are usually represented in F2 vs. F1 or F3 vs. F1 diagrams. This exploratory study was designed to investigate if vowel formants could be clearly depicted in three dimensions of F1, F2, and F3 and if vowel space centroid values would provide new information about vowel spaces. Three-dimension vowel space plots of the English vowel space were compared to the commonly used 2-dimension F2 vs. F1 trapezoid /I, u, ɒ, a/ and the triangular /i, u, ɑ/ schematic diagrams.

Description of the investigation and theoretical context
An uttered speech sound is a wave function with a specific physical shape. This wave can be converted by a Fourier transform into the visual depiction of frequency. Formants F1, F2, and F3 can be created by binning the areas of frequency-intensity, but each formant is a component of the overall sound wave.

We created 3-dimension line plots that show the relationship of formants to each other in terms of frequency (Hz.) plots. Our interest in F3 is further justified by entropy evaluations of vowel formants which we have found in some cases to be influential in defining the sound.

Methodology
A female and a male speaker of English produced recordings of the words containing /iː, ɪ, e, æ, uː, ʌ, ɔ, ɑː/ (bit, beat, bet, bat, boot, butt, bought, and bot), which should represent the corner vowels and edges of the vowel space. Formants were determined and this data was imported to computer programs written specifically for this evaluation. Vowel formant values exclude the bounding consonants, /b, t/.

The first program used the formant values to compute the location of each vowel centroid in three-dimension formant space. Plots were made for all eight vowels and these plots were rotated in space to examine the vowel locations and relationships to each other.

A second program used the vowel values and to create a centroid value for each vowel and the entire vowel space, which was also plotted. A centroid is

the mathematical integration of all the values of the vowel and represents the vowel in its entirety.

Results and conclusions

Vowel spaces were irregularly shaped, and did not match either the trapezoid-shaped or the triangle-shaped vowel space depictions.

The ability to rotate sound plots and to easily change axes helps to consider the relationship of formants to one another.

The centroid for the female speaker was found to be near the central value /ʌ/ (butt) in all formant plot positions, but for the male speaker it is located spatially between /i/ and /e/.

Future research

The vowel spaces of continuous speech utterances could be constructed to show how the vowel shape appears in its entirety in three dimensions. It is possible to determine the exact center location of each speaker's vowel space and with this knowledge we predict that it may be possible to identify a novel description of vowel space based on the geometric center.

Nominal predicates in pronoun resolution: an EEG study

Ann Hermansson, Fredrik Heinat, Eva Klingvall
Lund University, Lund, Sweden

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Theories on discourse accessibility and givenness suggest that nominal discourse entities vary with regard to how likely they are to act as antecedents to anaphoric expressions. Many factors are thought to affect the accessibility of a nominal entity in the mental model, including its morphosyntactic form and grammatical role. Moreover, previous research has shown that reflexive pronouns can be illusorily licensed, if they are preceded by a nominal that matches them in terms of morphosyntactic or semantic features (e.g. number and gender). While some nominals that are unavailable as antecedents by syntactic criteria (cf. Binding Theory) have been studied with regard to discourse accessibility and pronoun resolution, as well as to what extent they give rise to agreement attraction effects, non-referential predicate nominals have not yet been studied in a pronoun-resolution paradigm.

The current study made use of an EEG experiment in order to investigate to what extent predicate noun phrases are considered potential antecedents during online pronoun resolution. This was indicated by whether they gave rise to agreement-attraction effects. The results showed that a predicate nominal antecedent yielded a significant P600 response, and also a late sustained anterior negativity (L-AN) response against a baseline, referential nominal condition. The P600 and L-AN responses were interpreted as indications of the predicate nominals being realised as poor antecedents to the anaphoric pronoun. However, gender congruency between the predicate noun phrase and the anaphoric pronoun attenuated the P600 amplitude. The amplitude attenuation was interpreted as the predicate nominal being falsely perceived as an antecedent to the anaphoric pronoun, as indicated by the presence of agreement attraction effects.

The results therefore suggest that a predicate noun phrase is considered a potential antecedent to an anaphoric pronoun during the early stages of pronoun resolution.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Testing causal primitives in modals

Angelica Hill
University of Massachusetts Amherst, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Research question
Assuming that both deontic modals and causatives share causal meaning (Ilič, 2013,2014), is it possible to prime the production of a deontic modal after producing a sentence containing a causative?

Design
Both online production experiments made use of the sentence recall task in English. Following Branigan & Pickering (2017), we assume that structural priming can be used to target shared meaning between two constructions. Exp.1(n=48) and 2 (n=46) were designed similarly. For each trial (n=24), participants were asked to read aloud and memorize a target sentence in the simple past (control) and subsequently, a prime sentence which contained one of three conditions: modal-had to, causative-made, control-simple past (Exp. 1), or one of two conditions: causative-made, control-simple past (Exp. 2). Afterward people were prompted to recall the target sentence given a memory cue (verb). Crucially, participants were asked to utter a sentence containing had to 36 out of 72 total trials, thus people were biased to insert had to during recall. Counterbalancing was done using the Latin Square method over 6 experimental lists, and which sentence would be asked to be recalled was randomized. A priming effect was determined based on the difference in % of had to insertion during the recall period given the type of prime sentence.

Results
In Exp. 1 people inserted had to after a modal prime 46.57% of the time, after a causative prime 35.71% of the time, and after a control prime 30.93% of the time. The rate of had to-insertion after uttering a modal prime was significantly higher than the rate of had to-insertion after uttering a control prime ($p = < .01$). Although we did not find a significant difference of insertion between the causative and control prime ($p = .10$), there was a 4.78% difference between the causative prime had to-insertion rate and the control prime had to-insertion rate, which is consistent with the ~5% priming effect rate of other priming studies. In Exp. 2 participants inserted had to after a causative prime 38.55% of the time and inserted had to after a control prime 33.87% of the time ($p = .086$), a difference of 4.68%. To compare the rate of had to-insertion for the causative versus control condition across the two experiments, we combined

the data and used similar analysis as Exp. 1 and 2. Across the two experiments the rate at which people inserted had to after a causative prime was significantly higher than after a control prime ($p = .0015$).

Conclusion

Sentences containing had to and made don't share identical syntactic structure. Thus, the shown priming effect suggests it may be a shared semantic representation prompting the two to be processed similarly, and given the semantics of made it's likely this shared meaning is causal in nature. The results of the present study provide evidence that had to and made share a semantic representation that is causal in nature, and that the currently independent formal analyses of the two may have to be reconsidered.

Glottal Stop Fades and New Tone arises: The Emergence of Tone 8 in Zhangzhou

Yishan Huang
The University of Sydney, Australia

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

This study explores what mechanism has triggered the emergence of a new tone in Zhangzhou Southern Min, a Sinitic dialect spoken in southern Fujian province of in China. The acoustic utterances from 21 native speakers suggests eight tonal contrasts in the synchronic speech of Zhangzhou, contradicting all auditorily-based documentations of a seven-way tonal system. The eighth tone emerges from those syllables that are diachronically transcribed with a glottal stop coda and assigned in Yangru tone in the Middle Chinese (MC) tonal category. However, synchronically, the coda is discovered undergoing deletion, leading associated syllables to become open and giving rise to a new tonal category (tone 8). The exploration questions the conventional way of relying on the preservation of MC tonal categories and the citation context to calculate the totality of tonal contrast in Sinitic languages. It also provides a new profile to the typology of speech variation and evolution in world's tonal languages.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Analyzing R deletion in Brazilian Portuguese

Ana Paula Huback[1], Raquel Márcia Fontes Martins[2]
[1]Columbia University, US
[2]Universidade Federal de Lavras, Brazil

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

This paper provides diachronic and synchronic evidence that R deletion in coda position in Brazilian Portuguese (BP) is being implemented through Lexical Diffusion (Wang; Chen, 1977; Oliveira, 1997). R deletion in BP can occur either in middle or final position: amor > amô-- 'love', perfume > pe—fume 'perfume'. Historical documents, such as the Appendix Probi (3rd century) and the literary works by Gil Vicente, a Portuguese playwright from the 15th century, show the first occurrences of R deletion in final position in verbs. Leite de Vasconcelos (1970) and Chaves de Mello (1971) also show examples of that. Oliveira (1997, p. 33) analyzes this process through the lenses of Lexical Diffusion and proposes that "all phonological changes are guided through Lexical Diffusion."

Our experiment focused on final R deletion in nouns. Our goal was to understand which factors were favoring R deletion or retention, and if the change was spreading through Lexical Diffusion. Interviews with 30 native speakers were conducted, and a total of 2,606 occurrences of final R in nouns was gathered. Statistical methods were used to analyze the data. The deletion rate was 12%, and factors such as stress and social class were relevant to the process. The data also showed evidence for Lexical Diffusion, as the change is spreading through the lexicon according to token frequency, with frequent words changing first (Bybee (1995)).

Our paper makes a connection with Historical Linguistics as it analyzes R deletion diachronically, but it also presents results of a synchronic experiment.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Connection between lexical frequency and second language prediction

Haerim Hwang[1], Kitaek Kim[2]
[1]The Chinese University of Hong Kong, China
[2]Seoul National University, Republic of Korea

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Objective of Study: While first language speakers (L1) consistently demonstrate the ability to anticipate upcoming information based on linguistic cues, the findings for second language (L2) learners have been mixed, with some studies indicating they can, while others suggest they cannot. The current study explores the possibility that the inconsistent findings in L2 studies are attributable to similar psycholinguistic factors that can explain individual variations in L1 speakers. In particular, this study focuses on the impact of lexical frequency on L2 predictive processing, as increased usage of lexical items improves their accessibility, thereby making lexical frequency a crucial factor in facilitating predictive processing. The specific research questions we raise in this study are:

RQ1: Do L2 learners of English use verbal information to predict upcoming objects?
RQ2: Does the frequency of lexical items affect the predictive processing patterns in L2 learners?

Methodology
Thirty-seven L1-Korean L2-English learners participated in a visual-world eye-tracking task where they listened to sentences while viewing visual scenes containing two objects. The task comprised 20 critical items (alongside 36 fillers) distributed in a 2 × 2 Latin square design, with the factors "verb type" and "frequency", as shown in (1)–(4).

(1) higher-frequency, restrictive: The doctor opened the blue box.
(2) higher-frequency, neutral: The doctor moved the blue box.
(3) lower-frequency, restrictive: The surgeon opened the blue box.
(4) lower-frequency, neutral: The surgeon moved the blue box.

The verb types included: semantically restrictive verbs that limited the possible object options between the two visual stimuli, such as "open" for the box (target) versus the chair (competitor); and neutral verbs that could be used with either object, such as "move" for both box and chair. Importantly, the frequency of the subject NP's head was either relatively high (e.g., "doctor") or low (e.g., "surgeon"). For our statistical analyses, we computed a target advantage score as a dependent measure by subtracting the fixation proportions of the competitor from that of the target in each 20-ms time bin. From the verb offset to the onset of the object NP head served the prediction period (e.g., "the blue").

Results

Analysis of target advantage scores during the prediction period using a generalized additive mixed model (GAMM) where "verb type" was included as an independent variable (for RQ1) showed no discernible differences between the restrictive and neutral conditions in L2 learners' prediction patterns. However, a subsequent GAMM incorporating both "verb type" and "frequency" as independent variables (for RQ2) revealed a significant modulating effect of "frequency": L2 learners directed their gaze towards the target object upon hearing the restrictive verbs exclusively under the higher-frequency condition.

Conclusion

While L2 learners retain the capacity to make predictions using linguistic cues, their ability can be influenced by the frequency of lexical items. Based on substantial research highlighting the impact of lexical frequency on L1 processing, this study's findings suggest that both L1 and L2 speakers share similar predictive processing mechanisms, which are influenced by the same psycholinguistic factors, such as lexical frequency.

Differences in morphosyntax between ADHD and autism students

Dimitra Katsarou, Asimina Angelidou
University of the Aegean, Rhodes, Greece

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

ADHD is one of the most frequent neurodevelopmental disorders that is often found in primary school students and includes the characteristics of inattention, impulsivity and hyperactivity that persist over time and cause significant difficulties in language development (Pseftogianni & Katsarou, 2022). Language development in children with ADHD is slow, it is not the only difficulty of children with ADHD, but it is considered an important condition for the emergence of more complex speech disorders and learning difficulties (DuPaul, et al., 2013).

At the same time, high-functioning autism is a serious neurodevelopmental disorder and according to the DSM-5 children who fall into this spectrum are characterized by deficits in language. Deficits in morphosyntax and pragmatics in relation to language development are among the biggest language problems in children with high-functioning autism (Rumpf, Kamp-Becker, Becker & Kauschkea, 2012).

The aim of the present research was to measure morphological skills of children with ADHD and high-functioning autism, matched for age and non-verbal intelligence. The assessment tool used is the LAMDA (Skaloubakas & Protopapas, 2007) which was administered to 15 students with ADHD and 15 students with high-functioning autism in the 2nd - 4th grade, who had assessment from KEDASI.

Parameters were tested such as: speed and accuracy of picture and word recognition, historical and grammatical spelling, oral and written comprehension, syntax such as sentence completion and analogies, attention span, non-verbal intelligence and musical skills. According to the results, difficulties are found in both spelling and syntax without particular differences between students with ADHD and students with high-functioning autism. The analysis of errors per item, however, provided useful information for the development of morphosyntax between the two groups.

References

DuPaul, G., Gormley, M.J., Laracy, S.D. 2013. Comorbidity of LD and ADHD: implications of DSM-5 for assessment and treatment. Journal Learning Disabilities, 46, 43–51

Skaloubakas, X., Protopapas, A. 2007. Learning Skills and Weakness Detection Software. LAMDA Grades B'-D Primary. Athens: YPEPTH- EPEAEK

Rumpf, A., L., Kamp-Becker, I., Becker, K., Kauschkea, C. 2012. Narrative competence and internal state language of children with Asperger Syndrome and ADHD. Research in Developmental Disabilities, 33(5), 1395-1407.

Pseftogianni, D., Katsarou, D. 2022. Assessment of learning disabilities and ADHD in the context of psychoeducational assessment. In M. Konstantinou, I. Papageorgi & A. Malegiannakis (Eds.), Applications of Psychometry in Education and Educational Psychology (pp. 210-243). Gutenberg Publications.

L1 and L2 processing of different word orders

Maria Kharchevnik[1], Natalia Slioussar[1,2]
[1]Saint-Petersburg University, Russian Federation
[2]HSE Moscow, Russian Federation

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Many studies focus on the processing of flexible word order, but mostly by native speakers. We compared how L1 readers of Russian and L2 learners (Chinese speakers) process two word orders: the canonical SVO and the non-canonical OVS, also manipulating the context factor. In the first experiment, target sentences were presented in isolation, while in the second one, two types of context sentences were used: presupposing 'given – new' ('appropriate') and 'new – given' ('inappropriate') order in the target sentences, like in Kaiser and Trueswell (2004) and in Slioussar (2011) studying Finnish and Russian L1 readers. Online and offline measures were used: word-by-word reading times, interpretation accuracy (assessed through comprehension questions) and, for native speakers only, acceptability ratings.

Both the syntax (word order) and the context factor played a role for both L1 and L2 readers, but in a different way. L2 participants had more problems with the non-canonical order (longer reading times, lower interpretation accuracy), while for L1 participants, there were no significant differences between the two orders in appropriate contexts. Without context or in inappropriate contexts, the OVS order received lower acceptability ratings and was read slower than the SVO (interpretation accuracy was always ceiling for native speakers). It was also evident that for native speakers, syntax and context had a global influence (e.g. they were sensitive to the fact that the canonical word order has fewer contextual restrictions than non-canonical ones), while for non-native readers, local properties of noun phrases play a more important role (their case, their position in the beginning or at the end of the sentence, whether they are given or new).

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Further research is necessary to find out which differences between L1 and L2 readers are universal and which are specific for Chinese speakers learning Russian and are due to the differences between Chinese and Russian word order.

Processing verse and prose: intonational differences

Maria Kokhova[1], Natalia Slioussar[1,2]
[1]Saint-Petersburg University, Russian Federation
[2]HSE Moscow, Russian Federation

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

It is debated in the literature whether poetry is characterized by specific intonational properties — see e.g. Skulacheva (2017) and Kostyuk (2017) for the discussion of Russian verse and prose. In the present study, we look at this question from the perceptual angle. Are verse and prose distinguishable on the basis of intonation alone? To find out, we used delexicalized fragments of Russian verse and prose. We conducted a behavioral and a neurophysiological experiment to reveal perception differences on both conscious and unconscious levels.

16 prose fragments were taken from several 19th and 20th century Russian novels. 16 poems were written during the same period and ranged from syllabo-tonic (dominant in the classic Russian poetry and still the most widespread today) and dolnik (less regular tonic meters) and to free verse.

In the behavioral experiment, we asked participants (n = 41) to listen to delexicalized fragments (as long as they would like to) and to determine whether they were poetic or prosaic and analyzed accuracy and reaction time. Recognition accuracy of classical verse and dolnik was significantly higher than that of prose and other types of poetry. Free verse was recognized the least successfully, which was statistically significant. As for reaction time, no statistically significant differences were observed. The results indicated that not only intonation is required to distinguish between verse and prose but also rhythm, pausation, etc.

In the EEG experiment, the participants (n = 19) were told that the study aimed to determine the impact of different types of intonation on their emotional state. After listening to each delexicalized fragment, the participant was asked to answer a question about their current emotional state (these data were not analyzed). A statistically significant increase of the gamma spectral power was found when listening to fragments of prose in comparison to poetry. The alpha spectral power was significantly higher when listening to poetic fragments than when listening to prose. When comparing the «listening to delexicalized prose» condition with the «listening to delexicalized poetry»

condition, differences were found in the ipsilateral temporal electrodes T4 and T5. Thus, this study revealed differences in the perception of poetic and prosaic intonation, both on conscious and on unconscious levels.

Analysis of mRNA-vaccine posts on Swedish Twitter data

Dimitrios Kokkinakis[1], Bastiaan Bruinsma[2], Mia-Marie Hammarlin[3]
[1]University of Gothenburg, Sweden
[2]Chalmers University of Technology, Sweden
[3]University of Lund, Sweden

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

The global mass vaccination against COVID-19 (caused by the severe acute respiratory syndrome coronavirus-2), using the mRNA technology launched in 2021, have been hampered since the beginning, by rumours and fear among vaccine critical individuals that these "new" vaccines might be more dangerous to the health than other, "traditional" ones. As vaccine development is a time-consuming process, taking years to complete, Moderna Biotechnology, Inc. delivered an mRNA vaccine (design completion, manufacturing, testing, evaluation, approval) in less than one year from the discovery of the novel coronavirus. This rapid development caused concerns around efficacy and safety. Moreover, during early 2022, a medical journal article [1] written by a group of medicine researchers at Lund University, Sweden, spawned an already established vaccine rumour that somehow the mRNA vaccine is getting into and altering the human DNA. Having as an empirical case the fast spreading of this article, we analysed Swedish-language tweets discussing mRNA vaccines posted in 2022 using a mixed methods sequential explanatory design (initial computational distant reading analysis based on structural topic modeling and sentiment analysis, followed by a close qualitative reading and thematic analysis of the results).

The aim of this study was to use Swedish social media data to capture public perspectives and sentiments regarding the abovementioned study on possible effect of the novel mRNA vaccines that became massively available to the public during late 2021. The intention is to understand the key issues (topics/themes) that have captured public attention in Sweden, as well as the barriers and facilitators to successful or not mRNA vaccines.

Twitter was searched for Swedish tweets related to "mRNA" and "vaccine" (n=2030) during 2022. An unsupervised machine learning approach (i.e., structural topic modeling, STM) was used to identify topics from these tweets, with each topic further grouped into themes using manually conducted thematic analysis. Sentiment analysis of the tweets was also performed using a Swedish version of the rule-based machine learning model VADER [2].

Furthermore, the sentiments and posted time of the tweets were used as covariates to the STM model.

The qualitative thematic analysis of the quantitative results of the STM (9 topics), revealed seven overarching themes that drove the interlocutors' distrust in mRNA vaccines, the most prominent three among them being the worry that mRNA vaccines would alter human DNA, the technology's alleged experimental nature, and the potential adverse side effects from these vaccines.

The rapid production time in mRNA vaccine design and development spotlights some of the more insightful concerns of the public. The findings show that the mRNA rumours are not primarily based on ignorance, but rather on distrust regarding the officially sanctioned, positive narrative of new vaccine technologies, expressed through what we term counter-scientific argumentation. The results of our study may facilitate the formulation of more rigorous strategies to improve novel vaccine uptake (e.g. mRNA). The results further highlight the key processes that require attention in vaccination planning and may provide some guidance for overcoming evolving barriers related to novel vaccines.

[1] https://doi.org/10.3390/cimb44030073
[2] https://doi.org/10.1609/icwsm.v8i1.14550

Decoding language dominance in Spanish-English bilinguals' code-switching acceptability

Bryan Koronkiewicz[1], Rodrigo Delgado[2]
[1]The University of Alabama, US
[2]University of Illinois Urbana-Champaign, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Language dominance, marked by observable asymmetries in a bilingual's language usage, is a pivotal area of research in bilingualism. Various factors shape language dominance, including age of acquisition, environment, attitudes, and identity (Grosjean, 2012). Understanding this complexity is essential for comprehending how bilinguals navigate their languages, including intra-sentential code-switching (CS). Bilinguals can switch effortlessly, even when dominant in one language (Pérez-Leroux et al., 2014). However, it is less clear how structural constraints in CS relate to language dominance. Some argue the dominant language dictates switching patterns (Petersen, 1988), while later research challenges this stance (Liceras et al., 2016). It is also unclear how these claims differ from issues of language proficiency. Although the dominant language often exhibits higher proficiency levels, language dominance and proficiency are distinct aspects of a bilingual's linguistic profile. Given that proficiency can impact switching patterns (Toribio, 2001), it is crucial to disentangle the two for a better understanding of their impact.

This study investigates language dominance regarding acceptability ratings in a Spanish-English CS judgment task among highly proficient adult early bilinguals (N=22). The task included stimuli (N=24) from grammatical and ungrammatical switch conditions (1), targeting well-documented constraints related to pronouns, auxiliary verbs, and negation. Language dominance was measured using the Bilingual Language Profile (BLP; Birdsong et al., 2012). If dominance resembles proficiency, there should be less distinction between conditions the more a participant is dominant in one language or the other.

(1) a. Hace un minuto yo pedí a beer at the bar.
Hace un minuto yo pedí a beer at the bar.
b. * Hace un minuto yo ordered a beer at the bar.
'A minute ago I ordered a beer at the bar.'

The results show that language dominance does not have an overall impact on the acceptability ratings. Regardless of whether a participant was closer to

the middle of the BLP scale (i.e., "balanced") or not, participants rated the grammatical switches consistently more acceptable than the ungrammatical switches. These results were parallel for all four subcomponents of the BLP as well, as there were consistent ratings based on dominance regarding language history, use, self-rated proficiency, and attitudes.

Overall, this study uncovers that, unlike language proficiency, language dominance does not significantly affect acceptability ratings in intra-sentential CS among highly proficient adult early bilinguals. This underscores the need to distinguish between the two, and it continues the line of work targeting nuanced investigations of linguistic variables in research pertaining to CS and bilingual language use more generally.

Non-intervocalic geminates in Malayalam

Arya KS
Christ University, India

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Most of the existing phonetic studies on word-medial geminates have investigated durational and non-durational correlates of intervocalic singletons and geminates (Lahiri and Hankamer 1988). Apart from intervocalic geminates, word-medial non-intervocalic geminates across languages can be pre-consonantal, post-consonantal or non-vowel adjacent geminates (Pajak 2013). Malayalam has post-consonantal word-medial (also called as post-sonorant geminates in Mohanan 1986) geminates, which are always preceded by liquids (see 1c). The words given in (1) are of Sanskritic-origin (Mohanan 1986), and those in (2) are of Dravidian-origin. Both are morpheme-internal non-intervocalic geminates.

(1) a. [arkkan] 'sun'
b. [karppu:ram] 'camphor'
c. [kambi] 'metal rod'

(2) a. [ku:rkkam] 'state of snoring'
b. ku:rkka] 'country potato'

McCrary (2004) analyzed non-intervocalic geminates in Italian and highlighted the need to examine the phonetic environment of geminates. This study echoes the question raised by Kawahara (2015),which was to find out if non-intervocalic geminates will also be instantiated by the phonetic correlates of intervocalic geminates. Previous studies on non-intervocalic geminates, have examined the perceptual salience of singleton/geminate contrast in Russian (Dmitrieva 2012, 2017) and Moroccan Arabic (Pajak 2009, 2013) taking into account the infrequent distribution of non-intervocalic geminates cross-linguistically.

A survey of 100 studies on the phonetic correlates of singletons and geminates were carried out, based on which durational correlates like consonantal duration, duration of previous vowel and following vowel and voice onset time were investigated, as these were examined the most, cross-

linguistically. Two production experiments were conducted in this study, wherein the first experiment investigated the consonantal duration of post-consonantal geminates formed in sub-compounds in Malayalam. In the second experiment, along with closure duration, voice onset time, duration of preceding and following vowel were examined in nonce words created with word-internal non-intervocalic geminates, which were preceded by rhotics, and were of an edit distance of 1-2 from lexical items in Malayalam. The sound stimuli were analysed and annotated using Praat (Boersma and Weenik 2017). A linear mixed effects model was fit into the data using R studio (2016). A post-hoc Tukey's test was carried out wherever necessary, using the multcomp package.

The closure duration of geminates was longer than that of intervocalic geminates. However, the overall duration ratio was smaller than what was reported by Local and Simpson (1999). Voice onset time also was longer in intervocalic and non-intervocalic geminates in the second experiment. The duration of the flanking vowels remained the same in the two environments. The closure duration of non-intervocalic geminates were comparable to those reported in Russian and Moroccan Arabic, and were in the duration range observerved by Ladefoged and Maddieson (1996). Voice onset time could have lengthened as a function of duration. There was no clear instantiation of inverse CV patterning in the flanking vowels.

Study of addressee agreement in select EIA languages

Satyam Kumar
Indian Institute of Technology, India

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

The primary objective of this research is to examine the unusual phenomenon of allocutivity or addressee agreement (henceforth AAP)[1] and a number of politeness strategies in three languages from Eastern Indo-Aryan (henceforth EIA) family, i.e., Thethi, a Maithili dialect, Magahi, and Angika. I demonstrate that the utilization of the addressee agreement system in these languages contributes to the fundamental and significant role it performs in daily interpersonal contact and honorificity between a speaker and a hearer.

The prominence of honorifics, especially through AAP, in these EIA languages can mostly be attributed to the hierarchical nature of their respective communities. In these social settings, individuals are required to exhibit reverence for those who are older, regardless of the slight age disparity. Moreover, it is crucial for individuals to exhibit respect towards younger individuals as well, especially when they are from in-laws. Examples (1) - (2) are some representative noble data of addressee agreement and (3) represents vocative (one of numerous politeness strategies) from Thethi and Angika:

1. ma:sʈər sa:heb kha:na: kha:b kərai he-kh-in/kh-un/kh-en [Thethi]

teacher tilte.H food eat.NMLZ PROG be.PRS-3H/3H-2NH/3H-2H

'The teacher is eating food' (When uttered neutrally vs to a 2NH or 2H)

2. wẽ cinema: dekh-əl(ə)-k-ai/kh-o/hõ [Angika]

3NH.SG movie see-PST-3NH/3H-2NH/H

'He/She(H/NH) saw me.' (When uttered neutrally vs to a 2NH or 2H)

3. a:-/e:-ho:, iɖ̪hər a:-bɔ: nɛ̃
[Vocative, Thethi] VOC.2H here come-2H NEG/AFF

'(Please) come here.'

Based on the studies of Japanese, Basque, and Korean (Oyharcabal 1993; Miyagawa 2012, 2017 and others), AAP has been looked at as a root clause phenomenon. It has been studied in a number of other languages from across the world as well: Verma (1991) and Alok (2021) for Magahi; Haddican (2018, 2020) for Galician and some Southern Basque varieties. Most of the arguments in these studies say that the AAP locus is the very highest phrase of the clausal spine, like the SAP (Miyagawa 2012, 2017) or the cP (Portner et al. 2019a). However, researchers like Haddican and Alok argue that it is lower in the clause. In this study, I provide evidence that the aforementioned languages exhibit addressee agreement in both matrix and subordinate finite clauses. Consequently, I also argue that it is not accurate to categorize this phenomenon solely as occurring in root clause and that there is a possible involvement of the 'Left Periphery' coordinates (Rizzi 1997) in its representation.

[1] Antonov (2015): It's the way a non-argumentative addressee is encoded in language in certain sociopragmatic and syntactic situations in some or all main clause predicates. This indicates that the constructions in which a particular marker or inflection speaks for a null addressee are examples of allocutive agreement.

References

Bonaparte, L-L. 1862. Langue basque et langues finnoises. Londaon: Stange ways & Walden

Miyagawa, S. 2017. Agreement beyond Phi. Cambridge, Mass: MIT press.

Subbarao, K.V., et al. 1991. Syntactic strategies and politeness phenomena. International Journal of the Sociology of Language. Volume 1991, pp. 35-53.

Narrative skills of preschool and first school age children with Developmental Language Disorder (DLD)

Garyfallia Lemoni, Georgia Andreou
University of Thessaly, Volos, Greece

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Developmental Language Disorder (DLD), previously known as Specific Language Impairment (SLI), is a common developmental disorder that affects both preschool and school aged children. The language characteristics of children with DLD vary considerably with the main challenge being the learning of language structures and morphosyntactic and lexical limitations. The discourse abilities of children with DLD have been examined in two different ways: (1) focusing on the information content and information structure in the discourse and (2) focusing on the language content in the discourse. The first approach is referred in the literature as macrostructure analysis whereas the second is referred as microstructure analysis. The two most widely used models of macrostructure are story grammar and high point analysis which focus on the key components of a story, the sequence of events, and the episodic structure of a story (Justice et al., 2006).

In this paper the narrative skills of monolingual children with DLD and children with typical development (TD) are examined at a macrostructure level, using the most common methods of evaluating narrative skills a) story production and b) retelling. Our hypotheses are: a) the narrative abilities of children with DLD are more affected at the level of macrostructure than those of children with TD b) both groups of children present a weaker performance in story production with pictures rather than in retelling with pictures. The sample consisted of ten pre-school and school age children (five with DLD and five with TD), 5 to 11 years old, matched on chronological age. The Multilingual Assessment Instrument for Narratives (MAIN) (Gagarina et al., 2019) was used in order to assess the children of both groups in story production and retelling with pictures. The results confirmed our initial hypothesis that the narrative abilities of children with DLD are more affected at the level of macrostructure than those of children with typical development. More specifically, DLD children scored lower than TD children in the sections of story structure, structural complexity and internal state terms. Also, our second hypothesis that both TD and DLD children scored higher in the tasks of retelling with pictures than in the task of story-telling was confirmed. The findings are consistent with previous research showing that the stories

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

containing the least story grammar information and the greatest extraneous information were produced when children viewed the pictures without hearing an oral version.

References

Justice, L.M., Bowles, R.P., Kaderavek, J.N., Ukrainetz, T.A., Eisenberg, S.L., Gillam, R.B. 2006. The index of narrative microstructure: A clinical tool for analyzing school age children's narrative performance. American Journal of Speech-Language Pathology, 15, 177–191.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U.,Walters, J. 2019. MAIN: Multilingual Assessment Instrument for Narratives - Revised. Materials for use. ZAS Papers in Linguistics, 63. Greek version. Translated and adapted by Tsimpli, I. M., Andreou, M. &Peristeri, E.

Comparing language in typical and atypical Emirati children

Alexandra Marquis, Dimitrios Ntelitheos
United Arab Emirates University, UAE

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Our study compares language in Emirati children with autism spectrum disorder (ASD), intellectual disability (ID) and typical development (TD). ASD is a neurodevelopmental disorder that challenges one's social communication and interaction skills (American Psychiatric Association, 2013), while ID involves intellectual functioning and adaptive functioning problems (American Psychiatric Association, 2013). ASD children differ from ID children (e.g., Hetzroni et al., 2019). Yet, both groups may present language delays and/or disorders (e.g., Marrus & Hall, 2017). To date, there exists only limited experimental data on language performance of Emirati children with ASD (e.g., Shaalan et al., 2021) or ID (e.g., AlMuhairi et al., 2023). While we expected to find performance differences and similarities between ASD and ID children, we predicted to find lower success rates in those groups compared to TD children.

The participants reported here are part of an ongoing investigation evaluating language in atypical Emirati children. We report preliminary data of 13 ASD children (3 F; M age = 7;10, range = 5;9 to 10;1 years) and 7 ID children (5 F; M age = 11;6, range = 9;3 to 13;4 years) that were compared to 43 TD children (10 F; M age = 6;7, range = 6;2-7;5 years). Children were evaluated in Arabic. The test evaluations included four oral language assessment tasks (see Marquis, 2021 for details) evaluating language perception (phonological awareness & speech perception) and language production (morphological awareness & phonological production). The tasks were administered in a span of two to four separate sessions depending on the participants' ability. The ASD and ID participants' data were compiled and their average scores for each task were analyzed in terms of cut-off scores of approximately 1.25 standard deviations below the means obtained from the TD children in Marquis (2021). Thus, children who obtained a score below the cut-off score failed a particular language task. Results show that, overall, ASD and ID children have more difficulties with the language tasks compared to TD children, despite the ASD and ID children being older on average. For instance, when only 5 to 21% of TD children failed one task or another, we see that 33.33 to 91.67% of ASD and 33.33 to 83.33% of ID children obtained a score below the cut-off score. These results indicate that ASD and ID have more difficulties with all the language tasks compared to TD children.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Moreover, phonological production, the easiest task for TD children, with only 5% failure, is the most difficult for the ASD and ID children with 91.67% and 83.33% failure respectively. Surprisingly, speech perception was the hardest task for TD children, with 21% of failure, but was the task with the least failure for the ASD and ID children who reached only 33.33% failure each.

The data reported here inform us about certain linguistic similarities between Emirati ASD and ID children, while revealing their differences when compared to TD children. Our ongoing investigation will provide additional data from more atypical children.

Prosodic scope of discourse markers in French

Philippe Martin
Université Paris Cité, France

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Discourse markers in French, such as en fait and du coup, have no syntactic relationship with units that precede or follow them, but are necessarily realized with prosodic contours that define prosodic dependency relationships with the stress groups that surround them, and therefore determine their position in the prosodic structure of the sentence. This positioning determines the range on which bears the marker in the sentence, as encoded by the speaker, and decoded by the listener.

While the different sequences of words qualified as discourse markers, such as en fait and du coup in French, can belong to various morphological and syntactic categories (conjunctions, adverbs, imperative forms, nominal and verbal syntagms, etc.), two properties, one syntactic the other prosodic, are determining factors to confer this label on them:

1. They form sequences devoid of explicit syntactical dependency relationships towards the units that precede or follow them (Dostie, 2019; Tutin, 2019).
2. They form one or more rarely several stress groups, maintaining prosodic dependencies towards the stress groups which follow or precede them (case of postnuclei). These relationships define their position in the prosodic structure and determine their scope in relation to other sentence units.

To explain this last criterion, we will recall a few elements of prosodic grammar in French, based on a prosodic structure defined as groupings in several levels of minimal prosodic units (accent phrases), sequences of words containing only one stressed syllable (excluding emphatic stress), in final position in French.

We can then assume that the range of a discourse marker in speech is determined by its position in the prosodic structure of the sentence, implying (or not) a grouping with other lexical units, which will be considered by the listener when access to the meaning of the sentence.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

References

Dostie, G. 2019. Paramètres pour définir et classer les phrases préfabriquées : La vengeance est un plat qui se mange froid. Bon appétit ! . . . Cahiers de lexicologie, 114, 27-62.

Tutin, A. 2019. Phrases préfabriquées des interactions : quelques observations sur le corpus CLAPI, Cahiers de lexicologie, 114, 63-92.

A sentence comprehension test with whistled Spanish experts

Julien Meyer[1], Vincent Rolland[2], Tini Socas[3], David Díaz Reyes[4]
[1]Université Grenoble Alpes, CNRS, GIPSA-Lab, Grenoble, France
[2]Université de Rennes, Rennes, France
[3]Asociación Cultural y de Investigación de lenguajes silbados Yo Silbo, Spain
[4]Universidad de Las Palmas de Gran Canaria, Spain

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Object

In whistled speech, people articulate words while whistling and thereby transform spoken utterances by simplifying them into whistled melodies. Whistled Spanish (Canary Islands) and whistled Greek (Antia, Evia Island) remain Spanish and Greek but transformed phonetically. One of the most striking aspects of this natural whistled transformation is that it remains intelligible to trained speakers, despite a highly reduced acoustic channel to convey meaning. However, only one perceptual study with systematic measures on whistled sentences has been published in 1970. Indeed, this traditional practice is endangered and hard to test with speakers of remote language communities.

Speakers of non-tonal languages (such as Spanish, Greek, Wayãpi…) transpose linguistic segments – typically vowels and consonants - into whistles, lending alternative insights into how the phonetic expression of phonemes can be drastically reduced without hindering cognitive reconstruction. Just like for spoken speech - but even more - diverse types of contexts help (topic, situation, talker-listener relation, listener's feedback, redundancy). But the question remains about how sentences with little context are recognized.

To address these challenges we joined forces with actors of language revitalization and built an original experimental setting to test highly trained whistlers on long and unexpected sentences. Our aims were the following: (i) to evaluate recognition rates at different levels of these sentences (words, syllables, vowels), (ii) to understand better confusion patterns at the light of acoustic measures from a high number of whistlers.

Methodology

The audio/video recorded experiment was included into the biggest annual whistled speech contest in Canary Islands. 39 sentences were built with the help

of three whistled Spanish experts, so as to ensure that they were completely new to participants and judged as 'long, difficult, with minimum context' (for example: "the gran mother will pay tonight the hat and the shirt"). 22 whistlers took part in the contest. Each whistler had one minute to transmit a sentence, selected by chance, to the partner situated 20m apart and also selected by chance. Repetitions were allowed as well as whistled feedback. Answers were finally written by participants.

Results

General correct recognition rates (words: 65,5%, syllables: 69,1%, vowels: 78,9%) are inverse correlated with sentence repetition numbers (particularly after second repetition). 23% of the participants fully understood the sentence. Regular patterns of confusion clarify the picture, most notably vowel-consonant confusions in the highest frequencies and a vocalic confusion matrix showing that /a/ and /e/ were the best recognized (83%), /u/ the worst (60%). In parallel, the acoustic distribution of whistles corresponding to vowel production was analyzed: a Generalized Linear Mixed Model analysis reveals significant differences in frequencies for different vowels across whistlers, except for /o/ and /u/; where freq(/i/)>freq(/e/)>freq(/a/)>freq(/o/)=freq(/u/).

Conclusions

Results confirm that whistled speech enables highly trained specialists to understand unexpected long sentences without much context. Relatively high general rates of recognition were found, given the difficulty of the task. These results also explain why traditional use of sentence repetition generally limits to 2, and corroborate several whistled speech teachers' observations.

Elaborating Russian spelling-correction algorithms with custom n-gram models

Mikhail Minin, Olga Mitrofanova
SPbSU, Russian Federation

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

The aim of the paper is to compare and improve the effectiveness of different approaches to the task of automatic spelling correction of Russian social media texts in terms of qualitative evaluation of results and algorithm complexity.

Typos are the simplest type of spelling errors. Apart from the obvious task of text correction when entering search queries in the information retrieval system and exchanging messages in social media, an important area of application of automatic spelling correction algorithms is the preparation of text corpora for NLP tasks, such as building language models or improving the quality of speech recognition.

The dataset used in the study includes the corpus of LiveJournal blog texts consisting of 2008 sentences [https://www.dialog-21.ru/media/3838/test_sample_testset.txt] and a reference corpus with manual correction [https://www.dialog-21.ru/media/3835/corr_sample_testset.txt]. For our experiments from the diversity of approaches we selected a) Norvig's algorithm, which belongs to the class of context-free error models using Levenshtein distance and a frequency dictionary, b) Symspell algorithm, which is also based on Levenshtein distance, but works with an extended dictionary. We proposed and tested the context-aware extension of the given approaches with a bigram language model. In addition to basic operations performed by Norvig's and Symspell algorithms, the bigram model chooses the closest correct pairs for each pair of adjacent words with possible typos.

Results of experiments are defined in terms of classification errors (error matrix and the related parameters accuracy, precision, recall and F1-measure). Results of SpellRuEval competition [https://www.dialog-21.ru/en/evaluation/2016/spelling_correction/] were used as a baseline.

Norvig's algorithm implemented in pyspellchecker library [https://pypi.org/project/pyspellchecker/] was tested with two Russian frequency dictionaries consisting of 160K and 1mln word forms correspondingly [https://opencorpora.org/?page=downloads]. As the dictionary size increases, recall improves significantly, which corresponds to the intuition that more correct words are now included in the dictionary and are not distorted by the algorithm.

Symspell algorithm implemented in SymSpellPy library [https://github.com/wolfgarbe/SymSpell] was tested with the same dictionaries. As Symspell has a rich potential in working with n-gram language models, it gains in operation speed as compared with Norvig's algorithm.

Further, Symspell algorithm was tested with the language model including 900K bigrams. The original unigram dictionary was kept and used to identify and correct words written with hyphens. Thus, due to the extension Symspell uses a combination of the error model and the language model. Results outperform SpellRuEval in accuracy (best 71% vs. our result 91%) and approach in F1-measure (71%), (best 76%, 2nd result 65% vs. our result 71%).

We obtained evidence that it is necessary to use n-gram models of sufficient size in spelling correction. The quality of spelling correction has a significant impact on tokenization and POS procedures. In information extraction tasks it is important to correctly identify words that are included in the output of keyword and named entity extraction algorithms. These are primarily nouns, verbs and adjectives. According to error analysis results, the proportion of errors in these classes decreases after processing by the algorithms.

Does phonological change in Bengali-verb guide communicative intent?

Monsija Mitra, Pijush Kanti Gayen, Shankha Sanyal, Samir Karmakar
Jadavpur University, Kolkata, India

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Illocutionary acts capture speaker's intention and the recognition of intention is necessary for a successful communication. Listeners use several linguistic and extralinguistic cues to understand communicative intent and the linguistic cues are generated by speakers. Our objective is to observe how phonological change affects communicative intent. In Bengali, minimal phonological contrasts in some verbs express change of tense. We hypothesized that verb forms in minimal-pairs also change the speaker's intention, and both pair of the verbs expresses order and request acts with present and future tense, respectively.

We have selected Bengali verbs in minimal-pairs with the 2nd person familiar (e.g. /ɰmi/, /ɰomra/) agreement to meet the particulars of our assumptions. All the selected verbs are in CVCV pattern (e.g. /lek^{h}o/ ~ /lik^{h}o/). We assign CategoryA and B to both pairs respectively. We conduct Likert Scale (LS) and online recognition of intention probes (Reaction Time) to record the degree of intention ascribed to the form of verb and the underlying psycholinguistic process involved in the recognition of that intention.

Observations from the LS result comparing each intention between categories show that mean scaling of order is higher (Mean=3.57) for CategoryA, request is higher in CategoryB (Mean=3.18), reminder is higher in CategoryB (Mean=3.09) and, suggestion is higher in CategoryA (Mean=3.27). Then, ANOVA was performed where we observe how scale varies over an interaction between Speech-Act and Category. We found a strong effect of Speech Act on the response ($p<0.0002$, $F=47.78$), but the effect of Category was found insignificant ($p=0.33$, $F=0.93$) unlike interaction with the Speech-Act. The interaction effect (Speech-Act: Category) was found highly significant ($p<0.000024$, $F=6.67$), suggesting effect of Speech-Act on the responses Scale that depends on the levels of Category, and vice versa. The t-test was performed on ANOVA results to compare responses between two levels of Category for specific Speech-Acts. Statistical significances are as order ($p=0.006$, $t=2.76$), request ($p=0.02$, $t=-2.32$), and reminder ($p=0.0003$, $t=-3.59$). Suggestion ($p=0.49$, $t=0.68$) and offer ($p=0.66$, $t=0.43$) were not significant.

RT experiment results demonstrated that order intention recognition time for CategoryA was 2.50s (Accuracy=71.4%), while request intention recognition time for CategoryB was 3.36s (Accuracy=70.69%).

The LS experiment result illustrates that order is the dominant intention followed by suggestion within CategoryA, whereas request is dominant, followed by reminder and suggestion for CategoryB. We speculate that the act of suggesting can be performed through ordering and requesting; but reminding can only be performed through requesting. This assumption is supported by the RT result, as order recognition time is faster than request. Slower processing time for request entails higher cognitive load to elicit reminder and suggestion, whereas faster RT for elicitation of order suggests less complex processing due to overlapping with suggestion only. Thus, results from both experiments prove our first hypothesis that phonological change not only entails change in tense but also change in intention. However, it rejects our second hypothesis that CategoryA only expresses order and CategoryB request, but each form of the verb possesses an overlapping of several intentions.

Andalusian vocalism: comparative analysis of the cities of Malaga and Granada

Belén Reyes Morente
Universidad de Málaga, Spain

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

The research presented here focuses on the comparative analysis of the phonetic realization of vowels in the two major areas of Andalusian Spanish. Studies on dialectology and geolinguistics (Alarcos Llorach, 1983; Alvar, 1996; Salvador, 1977; Mondéjar, 1991) traditionally divide two large dialect areas in Andalusia: Eastern Andalusian and Western Andalusian. One of the features that separate both areas is the different behavior of the vowels, which is significantly more open and advanced in the eastern area. Navarro Tomás (1939) related the more open articulation of vowels with the eliminatin of /-s/ in coda. The different pronunciation of vowels was the object of study by several researchers (Alonso, Vicente & de Zamora, 1950; Bengoechea, 2006; Hernández-Campoy & Trudgill, 2002; López Morales, 1984; Martínez Melgar, 1994; Villena Ponsoda, 1987).

This paper presents the results obtained in an experiment conducted by the researcher in which the acoustic analysis of the vowels /e, a, o/ pronounced by four women, two from Malaga and two from Granada, is carried out. The four women are over sixty years old, but they differ by level of education so that in each city there is one informant with university studies and another with primary studies. The experiment consisted of three tests: reading a text, describing images and reading short sentences; and it was carried out through WhatsApp video call and audio. In each of the tests, the informants had to pronounce several minimal pairs of singular and plural (eg: boot/boots). The main hypothesis is that the acoustic distance between the final vowels of the singular and plural words will be larger in the informants from Granada than in those from Malaga. For the acoustic analysis, the Praat program (Boersma & Weenik, 2022) was used, which allows analyzing different acoustic parameters, in the case of this study it focused on the height of the formants F1 and F2. The statistical test chosen to be carried out with the SPSS program was a univariate analysis that allowed us to observe the behavior of the vowels /e, a, o/ taking into account the social variables mentioned above.

The results seem to confirm the proposed hypothesis, since they reveal that in women from Granada the acoustic distance between the vowels of the word in the singular and in the plural is larger than in women from Malaga. However,

in Malaga it is observed that the vowels can present features very similar to those obtained in Granada.

References

Alarcos Llorach, E. 1983. Más sobre vocales andaluzas. In J. Fernández Sevilla, H. López Morales, J. A. de Molina, A. Quilis & G. Salvador (Eds.), Philologica hispaniensia: in honoren Manuel Alvar (Vol. 1), (pp. 49-56). Madrid: Gredos.

Alonso, D., Vicente, A. Z., & de Zamora, M. J. C. 1950. Vocales andaluzas. Contribución al estudio de la fonología peninsular. Nueva revista de filología hispánica, 4(3), 209-230.

Alvar, A. 1996. Manual de dialectología hispánica. El Español de España. Ar

Speaker and prosodic peculiarity classification in emotional speech

Neda Mousavi[1], Sven Grawunder[1,2]
[1]Martin Luther University Halle-Wittenberg, Germany
[2]Max Planck Institute for evolutionary Anthropology, Germany

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Rhythmic patterns in speech have been presented in the literature not only as potential markers for individual characteristics but also as cues for emotion recognition. In this study, we used the German emotional speech corpus (VMEmo, Batliner, et al., 2000) to investigate the relationship between rhythmic metrics, emotion recognition, and speaker variability. The emotional data in this corpus were elicited through an experimental method in which participants interact with machines in tasks such as making an appointment. The experiment was purposefully designed to elicit various emotional fluctuations in the speaker as a result of the machine's responses. However, emotional fluctuations were not labeled directly, rather the corpus tags mark the speaker's linguistic behavior during emotional expressions. With our focus on rhythmic patterns, we examined, in particular, the prosodic dimensions of this linguistic behavior and considered the prosodic tags of each utterance, including peculiarities such as pauses between words, strong contrastive stress, pauses between syllables, syllable lengthening, etc.

For data preparation aligned with our goals, we specifically extracted data from individual speakers from the VMEmo corpus, excluding machine waveforms to emphasize human interactions. Finally, segmenting the waveforms and adding new tiers of vowel and consonant intervals to the Praat text grids ensured the availability of relevant data for rhythm analysis. Subsequently, we performed an analysis of the rhythmic patterns in each utterance, including measurement of a number of metrics such as %V, ΔC, nPVIv, rPVIc, varcoC, and CV rate. In addition, we assessed PVI indicators in both the intensity and frequency domains, yielding a comprehensive set of 14 rhythmic features. We categorized the rhythm features into four different groups to allow a comprehensive comparison of their efficiency across time, intensity, and frequency domains: duration-based metrics, intensity-based metrics, frequency-based metrics, and a comprehensive group encompassing all metrics. This study is primarily concerned with two key questions: the accurate classification of prosodic peculiarities on the basis of rhythmic features and the differentiation of speakers on the basis of these same features.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

After using principal component analysis (PCA) for visual exploration of the variation within and between speakers, linear discriminant analysis (LDA) was used to detect linguistic behaviors and classify speakers based on rhythmic features. The results of LDA analysis are expressed primarily in terms of classification accuracy, which quantifies the performance of the model in classifying instances. In our study, these instances include various prosodic peculiarities and speakers and rely on rhythmic indicators for classification. Comparison of different sets of rhythmic features for prosodic peculiarity detection in our first phase of LDA yielded accuracies of about 0.40 for all metrics and about 0.36 for the duration-based, intensity-based, and frequency-based metrics, suggesting minor differences in their effectiveness. In the second phase of our analysis, the results show varying accuracies in speaker classification across rhythmic feature categories: 0.44 for all metrics, 0.19 for duration-based metrics, 0.17 for intensity-based metrics, and 0.18 for frequency-based metrics. These remarkable differences in accuracy between these feature categories suggest possible differences and need to be further explored.

Processing temporary syntactic ambiguities in Greek while reading

Michaela Nerantzini[1,2], Katerina Drakoulaki[1], Dimitris Katsimpokis[3], Antonia Boznou[1], Angeliki Andrikopoulou[1], Eleni Peristeri[4], Varlokosta Spyridoula[1]
[1]National and Kapodistrian University of Athens, Athens, Greece
[2]University of Ioannina, Greece
[3]Netherlands Comprehensive Cancer Organization (IKNL), The Netherlands
[4]Aristotle University of Thessaloniki, Greece
[1]National and Kapodistrian University of Athens, Greece

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Object of study
Syntactic ambiguity has been extensively researched in typical populations; however, there are very few studies focusing on Greek (Peristeri et al., 2020). The present study uses an eye-tracking methodology to investigate (a) how reading is affected in the context of temporary direct object/sentence complement ambiguity and (b) how morphological (case) cues are integrated in the processing of ambiguous sentences.

Methodology
60 Greek monolingual speakers were recruited (M = 27.25 years), with normal/corrected vision and no history of learning difficulties. Six lists of 38 sentences were constructed. Materials were designed to include subject/object ambiguities, in which (a) the argument structure of the embedded verb (optionally transitive vs. intransitive verbs) and the syntactic function of the DP following the embedded verb (object vs. subject) was taken into account, and (b) the ambiguity was resolved by morphological case. The following four conditions were included:

(1) Optionally transitive verb; object-subject reading; (grammatical) reading in agreement condition
"Kathos etroge/ ta revithia/ gemise/ me fuskales"
"Kathos etroge/ ta revithia/ gemisan/ me fuskales"

(2) Optionally transitive verb; object-subject reading; (grammatical) reading in case condition
"Kathos etroge/ tus lukumades/ gemise/ me fuskales"
"Kathos etroge/ o lukumas/ gemise/ me fuskales"

(3) Intransitive verb; subject grammatical – object ungrammatical reading in agreement condition

"Kathos etrehe/ ta revithia/ gemise/ me fuskales"
"Kathos etrehe/ ta revithia/ gemisan/ me fuskales"

(4) Intransitive verb; subject grammatical – object ungrammatical reading in case condition
"Kathos etrehe/ o lukumas/ gemise/ me fuskales"
"Kathos etrehe/ tus lukumades/ gemise/ me fuskales"

Results

Discrepancies in reading measures between subject and object reading were attested, both on the disambiguating word, and at the end of sentence as spill-over effects. For the disambiguating word, first fixation duration was significantly longer for subject-reading versus object-reading in agreement condition whereas first fixation duration was significantly longer for object-reading sentences in the case condition. For the end of sentence, first fixation duration was significantly longer for object-reading sentences in agreement condition whereas for the case condition, first fixation duration was significantly longer for subject-reading sentences. Similarly, for the disambiguating word, total reading time in the agreement and case condition was significantly longer for subject-reading.

Conclusions

Like previous findings (as manifested in self-paced reading and grammaticality judgment measures), our data suggest that participants successfully integrate S-V agreement and case marking information to disambiguate garden path sentences. As it is evident, when encountering the main verb, participants access the verb's argument structure, making predictions about the upcoming sentence constituents, contra to parsing principles such as Late Closure. A closer look into the regressions and spill-over effects will be carried out to investigate different aspects of the reading process.

The wide scope of German topical singular indefinites

Kim Tien Nguyen
Goethe-University Frankfurt, Germany

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Object of study

For German – a language with canonical S > O order that allows object fronting, there can be various constraints on the availability of wide scope in doubly-quantified sentences, with information structure being one of them. Endriss (2009) postulates that indefinites marked as aboutness topics take wide scope. Two topic marking devices in German are left dislocation and intonational marking (through a rising accent on the indefinite determiner). Notably, using the latter can also make a contrastive reading with indefinite narrow scope (more) available. So far, to the best of my knowledge, there has been no study that explicitly tests Endriss' (2009) hypothesis on the wide scope of topical (singular) indefinites in both canonical and noncanonical sentence structures. This is what the present study attempts to do in order to investigate the pragmatic influence of information structure and prosody on scope interpretation.

Methodology

The conducted web-based experiment has a 3x2 factorial design with the factors sentence structure (3 levels: SVO, OVS, and LD – left dislocation) and intonation (topic marking intonation with a rising accent on the singular indefinite determiner of the object quantifier, 2 levels: without, with). The quantifier fast jed- 'almost every' was used for the subject of target sentences and the ∃ quantifier ein- 'a/one' for the object. The experiment employed a sentence continuation task. Participants were first presented with a target sentence as audio. They then had to choose between three sentences (presented as text) to continue the target sentence they had just heard. These three continuation options allow the researcher to clearly identify which of the three possible readings (indefinite narrow scope vs. indefinite narrow scope, contrastive reading vs. indefinite wide scope) was interpreted when participants chose a certain continuation option. There were 24 target items, 48 fillers, and 24 participants (native German speakers).

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Results

The choice of the continuation option indicating indefinite wide scope reading significantly increased in the conditions where the indefinite object quantifier was marked as aboutness topic of the target sentence by means of left dislocation and/or intonational marking (SVO + with, OVS + with, LD + without, and LD + with). This increase can also be observed in the OVS + without condition, implying that this sentence structure can also serve to mark the indefinite object quantifier as aboutness topic. There were no significant interaction effects between the two factors. The more left-fronted the indefinite object quantifier was, the more available indefinite wide scope became, with or without topic marking intonation. A rising accent on the singular indefinite determiner of the object quantifier increased the availability of indefinite wide scope, regardless of sentence structure.

Conclusions

These findings indicate that sentence structure and intonation function independently from each other as topic marking devices that lead to topical singular indefinites taking wide scope, verifying Endriss' (2009) theory and the pragmatic influence of information structure and prosody on scope interpretation.

References

Endriss, C. 2009. Quantificational Topics: A Scopal Treatment of Exceptional Wide Scope Phenomena. Studies in Linguistics and Philosophy. Springer.

Toward a discription of Digor lexical prosody

Varvara Petrova
Lomonosov Moscow State University, Russian Federation

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Object of study

Digor is one of the two standartised forms of Ossetian language. It is spoken in Northern Ossetia, Russian Federation, by approx. 100,000 people.

While the stress in Iron dialect of Ossetian has received a detailed description (cf. Dzakhova, 2010; Borise & Erschler, 2023), no consensus on Digor prosody has been reached yet. Magomet Isaev described Digor lexical accent as dynamic, although claimed that intensity of vowels correlated with their phonological length [Isaev 1966]. Henderson, on the other hand, described a quantitative type of stress [Henderson 1949].

Given that both length and intensity are affected by vowel quality, the data discussed here consists only of utterances of disyllabic and trisyllabic words with the same quality of vowels within the word form. We compared the values for the two features, length and intensity, separately, expecting to see either a consistent dominance of vowel in the same position (fixed stress), a consistent dominance of a certain vowel within the utterances of the same word as opposed to a dominance of another vowel for another word (variable stress), or no consistent dominance of one vowel (stress does not correlate with a given feature).

Methodology

Data for this study were collected in Vladikavkaz in August 2023. Our consultants were aged 20 to 67, either university graduates or undergraduate students and bilingual in Russian. The utterances were recorded on Zoom H5 recorder using a WH20XLR headset microphone. The consultants were asked to translate the words selected from [Takazov 2003] from Digor to Russian. They were then asked to pronounce the correctly translated words in a natural manner and repeat them twice with a pause after each utterance. Thereafter the utterances of the same words were elicited in phrases like zʁaj […] "Say [word]" and čːi zɜqtajʔ – […] "What did you say? – [word]". The measurements were taken in Praat (Boersma & Weenink, 2023).

Results

Disyllabic words with "strong" vowels (/a/, /o/, /e/) express a tendency toward a higher intensity of the first vowel, while the mean values for "weak" vowels (/u/, /i/) of the first syllable show no difference compared to those of the second syllable. In trisyllabic words there is an overall tendency for the first vowel to be the most intensive one. While the majority of both disyllabic and trisyllabic words has the last vowel as a dominant one in terms of quantity, there are notable exceptions to this tendency. Those are words with the same length of vowels in contrast to the second vowel reaching 150% of the first on average among the words with similar vowel quality.

Conclusions

Our research has shown Digor stress to be variable and correlate with vowel length. This claim is supported by the instrospection of our consultants, which reported Digor words to rarely be stressed on the first syllable. The stress on the first vowel correlates with its length being equal to the second vowel rather than exceeding it significantly due to the general cross-linguistical tendency toward lengthening the word-final part.

Text format and poem processing: evidence from Russian

Tatiana Petrova, Elizaveta Puchkova
St Petersburg University, Russian Federation

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

This study aims to answer two questions. First, is there any effect of reading modality on poetic text processing? Second, what type of poem presentation contributes to more successful text processing? Fuchs et al. (2001) gave theoretical and experimental arguments for supposing that oral reading fluency may reflect overall reading competence. According to Hale et.al. (2007) reading aloud facilitates understanding of the text, despite the fact that it has a larger processing load. It was also shown that silent reading was stronger for retelling narratives (Schimmel & Ness 2017).

In a three-group experimental design, we investigated how readers process and comprehend information in a poem in three different formats: listening, silent reading of a written text, and reading a written text aloud. Russian native speakers (N=60, 28 f., aged 18-63 y.o., Mage=27±11 years) read (either silently or aloud) or listened to three different parts of the poem "Letter to General Z" by Joseph Brodsky. All stimuli were of the same size, topic and level of readability (checked via http://readability.io). The texts for the audio format were read by an announcer (male, native speaker of Russian, higher philological education, 53 y.o.). After reading (or listening) to the parts of the poem, participants were asked to answer factual and analytical questions, name keywords and estimate the subjective difficulty of each text (we used scales from 1 which was "very easy" to 5 – "very difficult").

The study was conducted in an open experimenter who could control that, indeed, when answering questions to the text, the participant does not return to reading this text, and when performing the "Reading aloud" task, he reads this fragment aloud, and not silently. The Mann-Whitney U test showed that listening a poem decreases the subjective difficulty of the text (p-value = 0.0075). Reading aloud is also a more resource-intensive process compared to listening to an audio recording (p-value = 0.0022). There are no significant differences in perception when reading aloud and silently. Data analysis of key-word structures allows to conclude that reading the poem aloud contributes to a more subjective interpretation of the content of the poem, focusing on the individual experience of the reader.

The overall results made it possible to rank the types of presentation of a poetic text depending on the task facing the recipient. Audio text is the easiest

format for the perception and interpretation of the content of the poem, according to the subjective opinions of the participants. The format "silent reading of a written text" makes one read more thoughtfully, as evidenced by the results of the analysis using the question-answer method. The most difficult presentation format, requiring the greatest cognitive costs from the recipient, demonstrating the individual differences of people when analyzing the content of the text, is the “reading a written text aloud” format.

Syntactic islands and focality in Russian

Ivan Rygaev
Institute for Information Transmission Problems, Russian Federation

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

One approach to syntactic islands appeals to information structure as an explanation. It states that islands result from a clash in information structure. Goldberg 2006 formulated it simply:

Backgrounded constructions are islands (BCI)

Ambridge & Goldberg 2008 tested WH-extraction out of the complement of three classes of verbs: light, manner-of-speaking and factive verbs together with their focality/backgroundedness status. As a metric of islandhood, they used the difference score between the acceptability judgements of sentences with and without extraction. As a focality metric, they used the negation test – to which degree negation of the main sentence implies the negation of the complement.

Their analysis shows strong correlation between the focus judgements and the difference score ($r = -0.83$, $p = 0.001$). The correlation between the focus and extraction acceptability judgements is lower, yet still significant ($r = 0.58$, $p < 0.05$). So the BCI hypothesis is supported.

In this study, I conducted a similar experiment for Russian. I used the same three classes of verbs, four verbs in each class:

Light verbs: govorit' (say), dumat' (think), schitat' (believe), polagat' (suppose)
Manner-of-speaking verbs: krichat' (shout), sheptat' (whisper), bormotat' (mumble), lepetat' (babble)
Factive verbs: pomnit' (remember), znat' (know), radovat'sja (be glad), sozhalet' (regret)

I tried to use the same difference score – a difference of acceptability judgements between a declarative sentence without extraction and an interrogative sentence with extraction:

Dina krichit, chto Seva ispachkal shtany (Dina shouts that Seva has soiled his pants)
Chto Dina krichit, chto Seva ispachkal? (What Dina shouts that Seva has soiled?)

Instead of the negation test, I asked informants to judge how natural the sentence sounds as an answer to the question to the complement of the main verb:

Chto Luba sshila? (What did Luba sew?)
Venya govorit, chto Luba sshila sarafan (Venya says that Luba sewed a sundress)

The experiment was conducted online using Ibex farm. It was advertised in social networks and attracted 515 participants.

Results show a significant correlation between the focus judgements and the extraction acceptability judgements ($r = 0.68$, $p = 0.0144$). The focus score is a good predictor of islandhood. Thus, the BCI hypothesis is supported.

On the other hand, it does not show a significant correlation between the focus judgements and the difference score. This is probably because the extraction out of 'chto' complements in Russian is in general more restricted than the extraction out of 'that' complements in English. The extraction acceptability judgements are near the lower boundary of the range (2-3 out of 7). In such case, the simple difference might not be a good metric to use.

Anyway, the study shows that the focality is not the only factor that plays a role in the islandhood status of the constituent. The focus judgements are about 2 points higher than the extraction acceptability judgements. Speakers would not always accept extraction out of the complement even when they accept the declarative statement as an answer to the question with the focus on the complement. Probably grammatical factors also play a role.

The perception of Mandarin affricates by English listeners

Yan Shi
University of Utah, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Second language (L2) listeners often experience difficulties in perceiving unfamiliar speech sounds, and the Perceptual Assimilation Model (PAM; Best, 1995) has been proposed to predict difficulties in perceiving L2 contrasts with respect to the mapping patterns. To explore L2 learners' perception of Mandarin consonants, previous studies have either used synthesized stimuli (Tsao et al. 2016) or within-language identification task (Wang & Chen, 2020). However, studies utilizing cross-language tasks only examined the mapping patterns of a subset of Mandarin consonants (Hao, 2012) or adopted a very low threshold (e.g., 25% in Wang & Chen, 2019). Nevertheless, no studies included a discrimination task to verify the predictions by comparing the perceptual assimilation and discrimination data. Thus, the present study aimed to directly examine the predictions of the PAM model and explore the perceptual performance of English listeners with no prior experience with Mandarin.

Sixty-three naive L1-English speakers were recruited in the study. Target stimuli consisted of five tokens of six CV syllables involving 6 Mandarin affricates (/tsi/, /tshi/, /tɕi/, /tɕhi/, /ʈʂi/, and /ʈʂhi/) bearing the identical falling tone. Nine stimulus pairs were created, differing in terms of either place of articulation or aspiration. Participants completed an AX discrimination task followed by a perceptual assimilation and goodness rating task with a 9-point Likert scale.

The results for the perceptual assimilation task showed that at the 50% threshold, participants mapped Mandarin /tsh/ and /ʈʂh/ onto English /ts/ and /tʃ/ respectively, resulting in the Two-Category (TC) pattern for /tsh/-/ʈʂh/. The remaining 4 affricates (/ts/, /tɕ/, /tɕh/, /ʈʂ/) were not categorized at this threshold. The Uncategorized-Categorized (UC) pattern was identified for /ts/-/tsh/, /tɕh/-/ʈʂh/, and /tsh/-/tɕh/ pairs; and the Uncategorized-Uncategorized (UU) pattern for /tɕ/-/tɕh/, /ʈʂ/-/ʈʂh/, /tɕ/-/ts/, /ts/-/ʈʂ/, and /tɕ/-/ʈʂ/. At a 70% threshold, the UU pattern was observed for all pairs.

The PAM model predicts that the easiest contrasts to discriminate will be those where two phonemes map to different native language categories (TC). The result proved that the TC pair was discriminated most accurately (99% correct). In addition, the PAM model predicts that for UC and UU pairs, there is a descending order of accuracy for Non-overlapping, Partial overlapping, and

completely overlapping sounds. Unexpectedly, the results of three UU-P pairs showed that /tɕh/-/tʂh/ and /tsh/-/tɕh/ were discriminated equally accurate (⩾98% correct) as the TC pair, but /ts/-/tsh/ was discriminated less accurately (60% correct), which was not predicted by the PAM model. Further investigation showed that those sounds in pairs with lower accuracy differed in aspiration while those with excellent accuracy differed in place of articulation.

In general, the PAM model made correct predictions on accuracy of discrimination for TC pairs and part of UU and UC pairs. The unexpected results might be due to L2 English listeners' relative sensitivity to place of articulation over aspiration that compensated for the narrow phonetic distance. The findings shed light on our understanding of the initial state of perception of Mandarin affricates and have implications on Mandarin pronunciation teaching.

Phonological errors in Swedish-speaking children with DLD

Simon Sundström, Charalampos Themistocleous
University of Oslo, Norway

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Introduction

Children with developmental language disorder (DLD) often exhibit deviant phonological development and are known to make more phonological production errors than their typically developing peers. Swedish-speaking children with DLD have been shown to perform below peers on both consonant and vowel measures, when assessed with a confrontation naming task, which is standard procedure in clinical speech and language therapy practice (Sundström et al., 2018). Errors can include, for example, substituting one sound for another or omitting sounds altogether. This study aims to quantify the phonological errors made by children with DLD compared to age-matched controls with typical development from South-East Sweden.

Methods

30 children with DLD and 29 age-matched typically developing children participated in the study. All children were between 4 and 6 years old, speakers of the Central Swedish dialect, without hearing impairments or known neurodevelopmental or psychiatric disorders. Children with DLD received a DLD diagnosis by a speech and language therapist and had mainly phonological and grammatical, excluding pronounced lexical, semantic, and pragmatic deficits. Phonological production was elicited with a 72-item confrontation naming task. A method that automatically calculates a composite score of phonological errors from written speech productions was used. This process involves converting the written speech productions to the International Phonetic Alphabet (IPA) and then calculating a composite score of the phonological errors produced by the speakers. These errors include deletions, insertions, substitutions, and transpositions and are calculated using the Normalized Damerau–Levenshtein Distance (Themistocleous, 2023).

Results

Using data from children with DLD and controls, this paper exemplifies the methodological approach and demonstrates that the composite score can objectively measure the effects of treatment on speech productions. The results

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

showed a significant difference in the phonological error scores between children with DLD and controls.

Discussion

The method employed in this study provides an objective and automated way to quantify the phonological errors often seen in children with DLD. It could be used to measure the effects of treatment on speech production. This approach could be valuable for clinicians and researchers working with children with DLD or other speech and language disorders.

Conclusion

Children with DLD exhibit a higher phonological error rate than their typically developing counterparts. This study employed a novel methodological approach to objectively measure these errors and the effects of treatment on speech production. The results demonstrate the efficacy of this approach and suggest that it could be a valuable tool for clinicians and researchers working with children with DLD or other speech and language disorders.

References

Sundström, S., Löfkvist, U., Lyxell, B., Samuelsson, C. 2018. Phonological and grammatical production in children with developmental language disorder and children with hearing impairment. Child Language Teaching and Therapy, 34(3), 289-302.

Themistocleous, Ch. 2023. Computational Language Assessment: Open Brain AI. arXiv, 2306.06693, 1-17. https://doi.org/10.48550/arXiv.2306.06693

Whistled phoneme categorization: the Vowel Space Range Effect

Anaïs Tran Tran Ngoc[1], Julien Meyer[2], Fanny Meunier[3]
[1]Cote d'Azur University, France
[2]CNRS, Grenoble, France
[3]CNRS, Nice, France

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Whistled speech is a natural speech form used for long distance communication in geographic conditions characterized by harsh topography and/or dense vegetation. It transposes spoken speech into whistles produced in the front oral cavity of the mouth. In non-tonal languages, this transposition corresponds to a whistled transformation of some of the characteristics present in spoken formants: the vowels are emitted at different whistled pitch levels depending on the frequency distribution of spoken vowel qualities. Indeed, whistlers approximate the spoken articulatory movements to pronounce phonemes while whistling. For example, in Spanish, whistled /i/ has the highest mean values of pitch, /e/ is lower, /a/ is even lower, and /o/ is the lowest. While whistled speech is not directly comprehensible to naive listeners, i.e. listeners that never heard whistled speech before, previous studies have proved they categorize whistled vowels correctly much better than chance.

In the present experiment, we used this natural, yet modified, speech form as a tool to investigate perceptual processes in language, more specifically the impact of intra- and inter- production variations. Previous experiments have shown that the mean rates of correct answers varied largely as a function of the vowels: with /i/ being better recognized than /o/, which is better recognized than /a/ and /e/. Here, we sought to compare the impact of whistled productions of two different Spanish whistlers, one having a larger whistled vowel space than the other. Our aims were the following: (i) to test the robustness of the vowel categorization hierarchy previously observed, (ii) to evaluate the impact of whistled vowel frequency range by comparing perceptual results based on the productions of two whistlers, (iii) and to test the possibility of a learning effect within the experiment.

We focused on four whistled vowels: /i, e, a, o/. We ran a behavioral experiment in which we asked 22 naive participants to categorize whistled vowel stimuli. The experiment was structured in 3 parts. In part 1, participants listened to 48 whistled vowels, corresponding to 12 versions of each vowel type. Part 2 was a training session with feedback with 16 stimuli. Part 3 was

similar to part 1. Stimuli were presented in a random order in each part. Two versions of the experiment were built with the productions of the two whistlers.

The results confirm that the categorization rate of whistled vowels is well above chance. A Generalized Linear Mixed Model analysis reveals significant differences in performance for the different vowels (/i/>/o/>/a/=/e/) and an effect of the whistler, where the larger vocalic range led to improved categorization. A learning effect was observed only for one vowel and for one whistler.

This experiment shows that the range of the vowel space used by different whistlers affects vowel categorization, but only suggests the possibility of a learning effect. These results also highlight the robustness of the vowel recognition hierarchy previously observed, and reflect a certain stability in the speech perception process when faced with inter-talker variability in a highly modified speech form.

The Russian dative of possession: an empirical examination

Eleanor Sand
University of Wisconsin-Madison, Madison, USA

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

An examination of the literature reveals a lack of empirical data-driven studies on the Russian dative of possession (Levine, 1990; Pete, 1979; and others). Reliance on informal acceptability judgements and small sample sizes has enabled the opportunity for flawed data and bias. I propose to fill this gap in the research on the Russian dative of possession.

This study focuses on inferences associated with the dative possessive constructions of kinship terms (KT) and alienable possession (AP), for which the speaker must decide between a dative case-marked personal pronoun (e.g., mne; "me") (hereafter referred to as dative pronoun) or possessive pronoun (e.g., moĭ; "my") for possessor. If the speaker selects the dative pronoun, the listener interprets this to emphasize the possessor's affectedness (Levine, 1990). According to the literature on KT, the conversational implicature that the possessum functions as, but is not the actual kinship relation arises from a dative pronoun (e.g., On byl mne-DAT/moim-INS o͡tsom; "He was my father"). For AP, a dative pronoun generates the inference that the possessum is being worn by the possessor (e.g., Ona razbila emu-DAT/ego-GEN ochki; "She broke his glasses") (Levine, 1990).

The experiment manipulated possessor type (dative vs. possessive) and context (KT: kinship vs. no kinship; AP: affectedness vs. no affectedness) in a 2x2 factorial design. Native Russian speakers (n=26) were asked to complete a contextualized forced-choice judgement task (12 KT & 12 AP targets, 24 fillers), reading a paragraph and choosing the most natural of two sentences as a continuation.

Logistic (binomial) regression was conducted on raw binary scores (fixed factors: possessor type, context; random factors: participant, trial). A pairwise comparison (emmeans) with the Bonferroni adjustment was conducted as well. KT results found no possessor type preference with kinship (estimate=-0.308, SE=0.227, z-ratio=-1.357, p=1.0000) and dative pronoun preference with non-kinship (estimate=1.744, SE=0.248, z-ratio=7.022, p<0.0001) (see Table 1). AP results showed no possessor type preference with affectedness (estimate=0.308, SE=0.227, z-ratio=1,357, p=1.0000) and possessive pronoun preference with non-affectedness (estimate=-1.503, SE=0.243, z-ratio=-6.194, p<0.0001) (see Table 2).

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Possessor type preference concorded with the literature on KT, however, AP results challenged the literature. Discrepancies between the AP results and the literature's claims can be explained by change in the Russian language over time, differences in formal vs. informal registers, or flawed data in the literature. Regardless, this study underscores the importance of experimental linguistics research and provides a fresh perspective on the Russian dative of possession.

Table 1. KT acceptability judgements (literature, results).
Possessive, kinship: Acceptable, acceptable
Possessive, non-kinship: Unacceptable, unacceptable
Dative, kinship: Acceptable, acceptable
Dative, non-kinship: Acceptable, acceptable

Table 2. AP acceptability judgements (literature, results).
Possessive, non-affectedness: Acceptable, acceptable
Possessive, affectedness: Unacceptable, acceptable
Dative, non-affectedness: Acceptable, unacceptable
Dative, affectedness: Acceptable, acceptable

References

Levine, J. S. 1990. Pragmatic implicatures and case: The Russian dative revisited. Russian Language Journal, 44(147/149), 9-27.

Pete, I. 1979. Datel′nyĭ pritiа͡zhatel′nyĭ v russkom ia͡zyke i ego sinonimy [The dative possessive in the Russian language and its synonyms]. Die Welt der Slaven, 24, 418-429.

Phonologic and orthographic routes in word processing

Natalia Slioussar[1,2], Daria Chernova[2]
[1]HSE Moscow, Russian Federation
[2]Saint-Petersburg University, Russian Federation

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

There is an important debate in psycholinguistics concerning visual word processing. Some theories argue for two independent access options, relying on phonological and orthographic representations. Other models claim that phonological representation is always activated during lexical access.

We provide a novel piece of evidence in favor of the latter approach. Unlike previous studies focusing on homophonous lexemes, we used homophonous and homographic word forms. This allows avoiding certain problems (e.g. different lexemes have different semantic requirements that are hardly possible to control for).

We ran two word-by-word self-paced-reading experiments. Russian nouns are inflected for six cases and two numbers, and certain forms in some paradigms are identical (syncretic), homophonous or homographic. We showed that homophonous forms create mild grammaticality illusions, while homographic forms do not.

In Experiment 1 (n=80), we compared sentences with non-homophonous (1) and homophonous (2) case errors. In 1st declension nouns, Gen.Sg forms can end in -y or -i, depending on the stem-final consonant. The latter affix gives rise to homophony with dative and locative singular. On the target noun, RT differences between correct and incorrect sentences were significantly smaller in the homophonous condition. On the following word, only the grammaticality factor was significant. Thus, the phonological route is activated at early processing stages, but orthographical and phonological representations are matched later.

(1) Otrytka dlya sestry/*sestre poterjalas' na pochte.
'postcard for sister.GEN/*DAT=LOC was-lost at post-office

(2) Kiosk u ostanovki/*ostanovke zakryt vsju nedelju.
'newsstand near bus-stop.GEN/*DAT=LOC is-closed all week'

Experiment 2 (n=65) builds on earlier work on number agreement attraction in Russian. Slioussar (2018) demonstrated that readers almost do not notice (i.e. do not slow down on) number agreement errors with syncretic dependent noun forms like in (4a) and especially like (3b), while similar errors in (3a) and (4b) provoke massive reading time delays. We used similar stimulus sentences, as well as sentences like (5a-b) with homographic forms. We replicated the results from Slioussar (2018), but found no evidence for attraction (i.e. for reduced error-related RT delays) with homographic forms like in (5a). Thus, the same spelling does not trigger grammaticality illusions — only the same pronunciation does.

(3) a. Bilet na koncert byl/*byli…
'ticket for concert.ACC.SG≠NOM.PL was/*were'
b. Bilet na koncerty byl/*byli…
'ticket for concerts.ACC.PL=NOM.PL was/*were'

(4) a. Tkan' dlja jubki byla/*byli…
'tissue for skirt.GEN.SG=NOM.PL was/*were'
b. Tkan' dlja jubok byla/*byli…
'tissue for skirts.GEN.PL≠NOM.PL was/*were'

(5) a. Maz' dlja nogi byla/*byli…
'ointment for foot.GEN.SG was/*were'
b. Maz' dlja nog byla/*byli…
'ointment for feet.GEN.PL≠NOM.PL was/*were'

Large language models in Aphasia

Charalambos Themistocleous
University of Oslo, Norway

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Language assessment plays a vital role in diagnosing and treating individuals with speech, language, and communication disorders, which often arise from neurogenic conditions. These conditions can either be developmental or acquired. However, current assessment methods are manual, labor-intensive, and time-consuming to administer and score. Such methods may contribute to additional stress for patients, especially those already grappling with the difficulties associated with their conditions. As such, there is a pressing need for more efficient, streamlined approaches to language assessment. To address this need, we have developed Open Brain AI (https//openbrainai.com).

Open Brain AI is a computational platform designed to revolutionize language assessment. This platform harnesses the power of innovative AI techniques, including machine learning, natural language processing, large language models, and automatic speech-to-text transcription. With these technologies, Open Brain AI can automatically analyze both spoken and written speech productions across a multitude of languages. The platform is capable of assessing a wide range of linguistic features, providing a comprehensive overview of an individual's language abilities.

In this paper, we detail the development process of Open Brain AI and the various AI language processing modules that constitute the platform. We outline how machine learning algorithms and natural language processing techniques have been employed to facilitate the automatic analysis of speech and written text. Furthermore, we discuss the linguistic measurements obtained from discourse macro-structure and micro-structure, which serve as crucial indicators of language proficiency and potential areas of concern.

The implications of this platform extend beyond improved efficiency. Open Brain AI reduces the burden on clinicians by automating many of the tasks associated with language assessment, allowing them to dedicate more time and resources to direct patient care. This, in turn, may lead to improved outcomes for individuals with speech, language, and communication disorders, as they can receive more personalized attention and targeted interventions.

Open Brain AI is freely accessible to clinicians, researchers, and other professionals involved in language assessment and intervention. This open access nature of the platform democratizes the field, enabling clinicians and researchers from diverse backgrounds and settings to conduct critical data analyses and contribute to the collective knowledge base. As a result, Open

Brain AI holds the potential to drive further advancements in the diagnosis and treatment of speech, language, and communication disorders.

Additionally, the platform's multilingual capabilities make it especially valuable in diverse linguistic contexts. By providing automated analyses of speech and written text in multiple languages and dialects, Open Brain AI can address the needs of non-native speakers and individuals who speak multiple languages. This is particularly important given the increasing linguistic diversity observed in many countries.

By leveraging advanced AI technologies, the platform automates many of the tasks traditionally performed manually by clinicians, streamlining the assessment process and reducing the workload for healthcare professionals. Its freely accessible nature and multilingual capabilities make it an invaluable tool for improving the diagnosis and treatment of speech, language, and communication disorders across diverse linguistic contexts.

Prediction in L2 perception of reduced multi-word sequences

David Tizón-Couto[1], David Lorenz[2]
[1]Universidade de Vigo, Spain
[2]Lunds Universitet, Sweden

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

The cognitive entrenchment of frequently occurring sequences comes in the form of 'chunking' (accessing the sequence as a single unit) and of 'procedure strengthening' (predicting the next step in a sequence). Existing research attests to the effects of frequency and entrenchment of multi-word sequences in the native language, which we learn and shape continuously and intuitively (cf. Blumenthal-Dramé 2018). But how do they affect L2 speakers, whose acquisition of linguistic structures is top-down (through language teaching) but who might nonetheless also learn through usage? (cf. Ellis 2013; Supasiraprapa 2019).

The present study addresses the issue of receptive processing of multi-word sequences by means of a word-monitoring experiment with advanced learners of English. Recognition (response time and accuracy) of the element to in the construction V-to-Vinf was tested for full and reduced renderings ([tʊ] vs [ɾə]), conditioned by the general frequency of the V-to sequence and the transitional probability (TP) of to given the verb (V > to). The experiment has previously been carried out with native speakers (Authors 2019), so the results can be compared directly.

Results show that recognition is slower and less accurate with reduced items. This is mitigated when the sequence has a high surface frequency. TP has no such effect. Thus, advanced learners seem to profit from frequency-based expectations in speech perception when the input is reduced. Native speakers show a different pattern when reduction is present, most notably a chunking effect of high-frequency strings and a facilitating effect of TP.

We conclude that, firstly, (advanced) learners' access to linguistic structures is more compositional than native speakers'. Secondly, they do take recourse to entrenched sequences to recover reduced input forms, but do not seem to derive expectations from transitional probabilities; as TP is more complex than surface frequency, it probably requires more and richer usage experience to become part of a language user's intuitive perception strategies. Overall, reduction implies a greater setback for advanced learners, who do not (yet) rely as heavily on the statistical information and compensation strategies generally available to NSs (cf. Ernestus et al. 2002; Pickering and Garrod 2007). We

attribute these findings to learners' lesser experience with spontaneous speech and phonetic reduction.

References

Blumenthal-Dramé, Alice. 2018. Entrenchment from a psycholinguistic and neurolinguistic perspective. In Hans-Jörg Schmid (ed.), Entrenchment and the psychology of language learning, 129–152. Berlin: Mouton de Gruyter.

Ellis, Nick C. 2013. Second language acquisition. In Graeme Trousdale & Thomas Hoffmann (eds.), Oxford handbook of construction grammar, 365–378. Oxford: Oxford University Press.

Ernestus, Mirjam, R., Harald Baayen & Rob Schreuder. 2002. The recognition of reduced word forms. Brain and Language 81. 162–173.

Pickering, Martin J. & Simon Garrod. 2007. Do people use language production to make predictions during comprehension? Trends in Cognitive Sciences 11(3). 105–110.

Supasiraprapa, Sarut. 2019. Frequency effects on first and second language compositional phrase comprehension and production. Applied Psycholinguistics 40. 987-1017.

English tough-constructions and their analogues in Russian

Alina Tsikulina, Efstathia Soroli
University of Lille, France

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Evaluative constructions involving tough predicates (e.g., This book is easy to read) present atypical form-to-meaning mappings and vary across the languages of the world. In some languages (e.g., English) speakers use so-called tough-constructions (TCs) to evaluate an event/process – constructions in which the syntactic subject NP of the matrix sentence is logically the missed object of the embedded non-finite verb. In other languages (e.g., Russian), such a construction is not possible but speakers express evaluation with a variety of functional analogues: e.g., use of passives – constructions with action predicates realized with a reflexive coupled with a manner adverb; use of deverbals – constructions involving a deverbal noun realized as a restrictive prepositional complement; use of predicatives – constructions involving a topicalized NP in the accusative case, a predicative adverbial and a dependent infinitive. Despite a growing interest in TCs with respect to their syntactic-semantic unalignments and their high crosslinguistic variability, the inherent semantic and morphosyntactic properties of evaluative constructions, especially in languages such as Russian that do not offer tough movement per se, have been only superficially discussed. The aim of the present study is (a) to explore English TCs and their Russian equivalents; and (b) to identify the properties and the contexts of occurrence of the most frequent analogues coupling two types of data: corpus-based and experimental.

The corpus investigation of the present study was based on the Opus-corpus subtitles database and allowed to compare English TCs as source patterns with Russian parallel translations as target alignments. The experimental study consisted in collecting written data from 25 L1 Russian speakers for comparison with the corpus-based findings. More specifically, the experimental data were gathered using a written description task. Participants had to type in a short sentence to describe a set of pictures depicting different scenes using three target words: an animate or an inaminate noun, a tough adverb and an infinitive, counterbalancing their order of presentation.

The translation corpus findings suggest that Russian offers mainly constructions involving an impersonal predicative, alternatively extraposed and compact constructions especially when the source TC involved high transitive verbs, with an inanimate NPs coupled with single-scope adjective, or with an

animate NP coupled with a double-scope adjective. The experimental study confirmed the general preference of speakers for impersonal predicative constructions but also revealed other alternative constructions that do not appear in the corpus study (e.g., subject-reading, modal uses) which occurred especially when the target words were presented in tough-adverb+verb+NP order and when the scene involved an inanimate NP. Passive and Deverbal constructions, although described as frequent functional analogues for Russian in theoretical papers, only marginally occurred in the corpus and the experimental studies.

This work reveals several analogue evaluative constructions (extraposed and compact alternatives in the corpus study, and modal and subject-readings in the experimental study), never discussed before for Russian. These findings only partially support previous theoretical classifications in the domain of TCs and suggests that parallel corpus and experimental investigations are necessary for a fullest exploration of the grammar-thought relationship.

Experimentally comparing the learnability of rule interactions

Yuxuan (Melody) Wang
Harvard University, US
University College London, United Kingdom

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Many AGL studies appeared to test the learnability of different rule interactions. However, none directly showed Kiparsky's (1973) claim that opaque interactions are harder to learn with evidence of transparency preference, and many had methodological or statistical deficiencies. This project directly compares Feeding with Counterfeeding and tests whether opacity is less learnable - without teaching participants specific orderings but rather letting them choose the default order under a 'poverty-of-stimulus' paradigm - with a reasonable sample size and statistics. If transparency is preferred, significantly more participants should be disposed to learning forms derived from Feeding than Counterfeeding.

The two interacting rules were Vowel Harmony (/e/→[i], /o/→[u] within a word) and Palatalisation (/t, d/→[tʃ, dʒ] / _i). Participants learned them by observing how the plural and diminutive suffixes (-i and -a, counterbalanced) worked in an artificial language. Participants saw a picture of an object and the corresponding word stem, followed by another picture with either two or a smaller version of that object. Participants then listened to two suffixed forms, and chose the one they deemed correct. Feedback was provided immediately. Participants were expected to learn the phonological rules using the feedback, especially how the stem changed when -i was attached.

The experiment had three phases: Training, Verification, and Test. Training had two blocks for learning Harmony and two for Palatalisation. Verification was a two-block version of Training to confirm participants grasped the rule. Test phase had one Critical and three Filler blocks (Harmony, PalC1C2 and PalC3; counterbalanced). Critical trials created environments where participants could apply either rule first. Filler blocks aimed to check the participants' understanding of each rule. Thirty native American English speakers took part.

Results were analysed with mixed-effects logistic regression models implemented in R using lme4. The percentage of participants choosing the Feeding option for Critical trials was significantly lower than chance, and the accuracies in Filler blocks were significantly different from Critical trials, i.e.,

participants successfully learned two rules separately but refused to let them Feed. In short, this study disproves Kiparsky's idea that Feeding is preferred or the easier order to learn.

Besides the participants' lack of preference for Feeding, they also achieved higher accuracy when the correct option required no change from the UR in Training. These suggest a faithfulness bias, which could be explained by structural complexity: the more features are involved in a pattern, the harder it is to learn (Pater & Moreton, 2012). The grammar deriving Counterfeeding requires learning only vowel co-occurring restrictions (height), but learning Feeding requires knowing the constraints on alveolar stops/affricates (only affricates before [i]) besides vowel agreement. Participants' rejection to apply Palatalisation is also interesting. Since Palatalisation required subjects to generalise the rule from the final consonant to non-final ones, the participants could have (1) employed a more conservative learning strategy, or (2) believed non-final and final consonants are intrinsically different, like in NDEB. Becase the final consonant features a position theoretically more prone to changes, participants may naturally think only the final consonant could palatalise.

Nominal quantification and DOM in Northern Galilee Arabic

Aya Zarka[1,2], Aviya Hacohen[2]
[1]McMaster University, Canada
[2]Ben-Gurion University of the Negev, Israel

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Differential Object Marking (DOM) is a widely studied crosslinguistic phenomenon, whereby only certain sets of objects are overtly marked with case (e.g., Aissen, 2003; Bossong, 1991). While DOM has long been investigated in a variety of languages, research on Arabic DOM is relatively scarce. Several dialects of Arabic, including Lebanese, Syrian, and sub-dialects of Levantine Arabic, have been argued to exhibit a DOM construction (Aoun, 1999; Brustad, 2000; Khan, 1984). In Northern Galilee Arabic (NGA), unmarked objects lack overt accusative marking, while DOM marked objects carry the prepositional dative la- 'to'. This is illustrated in (1).

(1) a. ʃof-et s^{ʕ}-s^{ʕ}abeyy-e.
saw.1.sg the-young lady-f.sg
'I saw the young lady.'
b. ʃof-t-a la-s^{ʕ}-s^{ʕ}abeyy-e.
saw-1.sg-3.f.sg.cl DAT(DOM)-the-young lady-f.sg
'I saw the young lady.'

We present results of a gradable acceptability judgment task designed to systematically examine how the nominal properties of the direct object affect the distribution of NGA DOM. Building on previous theoretical works, we tested the individuation generalization proposed in Brustad (2000) and specifically, Zarka's (2021) proposal, that quantification is the key feature regulating NGA DOM. Departing from Zarka's assumption that quantification exclusively pertains to morphosyntactic countability, we identify the concept of perceptual discernability as an additional aspect of nominal quantification. We hypothesize that perceptual discernability, and not morphosyntactic countability, is the right dimension for characterizing the distribution of NGA DOM.

To test this hypothesis, we designed a gradable acceptability judgment task with three experimental conditions: count, mass, and – crucially – object mass nouns. The latter are nouns that pattern with mass nouns in terms of their morphosyntactic behavior, but unlike prototypical mass nouns, they refer to a

set of perceptually discernable individuals, and in this sense, they are similar to count nouns (cf. Barner & Snedeker 2005). As such, object mass nouns provide a perfect test case for whether NGA DOM is regulated by morphosyntactic countability or perceptual discernability. The experiment was conducted online, with the verbal stimuli presented as audio files. The manipulated variable included three DP-types: count, substance-mass, and object-mass. 48 adult NGA speakers provided judgments on a 6-point scale, with only the extremes of the scale explicitly labeled (1=btnḫkaš ('cannot be uttered'); 6=akid btnḫka ('can absolutely be uttered').

Our findings reveal that 'morphosyntactic countability' is not the right dimension for characterizing the distribution of NGA DOM. Instead, licensing of DOM depends on whether the direct object denotes perceptually discernible individuals. This analysis of DOM is novel, not only for Arabic but also when viewed crosslinguistically. We discuss some similarities between animacy and perceptual discernability, suggesting that both can be viewed as aspects of Prominence.

In sum, the results of the current study provide a first look at how noun type affects the distribution of NGA DOM. The range of nominal categories tested and the carefully controlled experimental manipulations yielded a set of acceptability data that is broader and more detailed than what has been previously discussed in the literature. Consequently, these findings allow for a more precise and more comprehensive mapping of the distribution of NGA DOM.

Sensitivity to classifier relation in a priming task

Jiahuan Zhang
The University of Hong Kong, Hong Kong, Hong Kong
The University of Cambridge, Cambridge, United Kingdom

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Chinese classifiers are unique linguistic units, encoding both grammatical and semantic information of their associated objects. Chinese speakers were often found to show sensitivity to classifier rules and semantics through the lens of cognitive processing tasks, which is the purported "classifier effect". However, it remains unclear whether the classifier effect is reflective of the linguistic sensitivity or universal conceptual structuring. To scrutinise this, a linguistically-implicit mediated priming task was administered to assess differences in sensitivity to information contained within classifier-related pairs. Stimulus trials included three components: 1) a prime that refers to a classifier's typical noun; 2) a mediator that denotes the major semantic feature of the corresponding classifier; and 3) a target corresponding to the prime.There are three conditions on stimuli pairing: grammatical pairs (G+S-) included classifier-matched prime and target but were semantically disassociated (e.g., tunnel-long- shorts); semantic pairs (G-S+) were semantically associated, without prime and target classifier match (e.g., rope-long-lipstick); and fillers (G-S-, e.g., leg-long-fan). Intra-group comparison showed that Chinese (n = 35, female=29, Mage=24.16, SD=2.58, age range:18-31) and English (n = 35, female=26, Mage: 22.31, SD=2.71, age range:19-28) L1 speaker groups both correctly produced more "yes" responses on semantic pair trials when being asked to decide if the prime and the target were associated with each other. Although both L1 groups performed comparably in responses under all task conditions, it only took Chinese speakers longer time to determine grammatical pairs than English speakers. Findings together support an "implicit classifier effect" from Chinese speakers' performance.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Perceptual adaptation to English-accented Mandarin Chinese

Kevin Yi Zhang
Department of Languages, Literatures, and Linguistics, US

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Previous research suggests that native English listeners can adapt to a foreign accent through exposure to multiple talkers with that accent. Adaptation tends to be dynamic, with the accuracy of transcribing noise-embedded accented speech increasing over time. While existing research predominantly focuses on English and emphasizes the importance of recognizing segments, the adaptation process has not been extensively studied in other typologically different languages. This study aims to investigate the role of segments and tones in native Mandarin listeners' adaptation to English-accented Mandarin.

Forty-nine native Mandarin listeners with little-to-no exposure to foreign-accented speech were recruited for a three-day training followed by a posttest. During training, listeners were randomly assigned to one of four conditions and transcribed 70 unique disyllabic words per day. In the three experimental conditions, listeners heard naturalistic (intact), monotonous (segmental), and tone-only (tonal) versions of the materials spoken by four English-accented talkers. In the control condition, the same stimuli were produced by four native Mandarin talkers. In the posttest, listeners were asked to transcribe either disyllabic or sentence-length materials (210 syllables) produced by a novel English-accented talker. The transcription materials were phonetically balanced and embedded in noise at a +3 SNR.

Binary accuracy data on segments, tones, and lexical items were analyzed using the glmer package in R. Throughout the three-day training, listeners in the intact condition were consistently more accurate than those in the segmental condition on segments ($\beta = 0.51$, $z = 3.69$, $p < 0.001$) and those in the tone condition on tones ($\beta = 1.59$, $z = 7.61$, $p < 0.001$). For within-condition comparisons, while listeners in the latter two conditions showed significant improvement on the second day and maintained their performance since, listeners' accuracy in the intact condition declined over those days, possibly due to fatigue and their high initial accuracy on the first day. In the posttest on disyllabic words, no significant differences were found among groups for either segmental or tonal accuracy. However, listeners in the control group performed significantly worse than those in the intact group when considering lexical accuracy ($\beta = -0.49$, $z = -2.12$, $p < 0.05$). For the posttest on

sentences, despite the absence of a global effect of condition, trial number positively predicted transcription accuracy for nearly all measures and in all conditions except the intact condition.

Taken together, since exclusive training on segments or tones did not provide listeners with an additional advantage in adapting to those specific aspects of accented speech, this study suggests the essential role of lexical information in guiding adaptation to low-level acoustic variations. Furthermore, the dynamic nature of adaptation was more evident in sentence transcription than in word transcription, likely because sentences offer more contextual cues that can help listeners gradually achieve higher and more stable recognition accuracy. The absence of significant differences between groups in the posttests could be attributed to the rapid adaptation process during testing itself and/or the considerable individual variability within the non-intact conditions. These aspects will be the focus of future research endeavors.

Processing Ambiguous Object Clefts in Mandarin Chinese

Jiayi Zhou, Rishabh Suresh, Vera Verrijt, Alessia Giulimondi
Institute for Language Sciences, Utrecht, Netherlands

https://doi.org/10.36505/ExLing-2023/14/0000/000000

Abstract

Research Questions Object clefts in Mandarin Chinese with ditransitive verbs such as gei ('give') are ambiguous as the clefted determiner phrase can, given the right context, be interpreted as either the direct object (DO) or indirect object (IO). As for the Mandarin sentence "puppy COP rabbit gave kitten DE animal", there are two readings: the DO-cleft reading is The puppy is what the rabbit gave to the kitten, and the IO-cleft reading is The puppy is who the rabbit gave the kitten to.

This study sought an answer to whether speakers have a preference for analysing an ambiguous object cleft as an IO or a DO cleft. Based on the 'Noun Phrase Accessibility Hierarchy' (Keenan & Comrie, 1977) in language production, IO-clefts were expected to be more difficult to access than DO-clefts. Furthermore, it investigated whether their syntactic preferences on the structure were modulated by semantic cues; in particular, the manipulation of animacy. Since the interactive model of parsing LAST (Townsend & Bever, 2001) suggests that the parser considers semantic cues, and the prototypical animacy in ditransitive structure is animate IO and inanimate DO (Buckle, 2017; Velnić, 2019), it was expected that within IO-clefts, those with an animate DO would be more difficult to process than those with an inanimate DO.

Methodology and Results A self-paced reading test with a moving window was adopted to test our hypotheses with a sample of 44 adult Mandarin native speakers. There were 128 stimuli divided into 4 lists of 32 stimuli each using the Latin Square design. Four conditions of stimuli are IO-cleft with animate DO, IO-cleft with inanimate DO, DO-cleft with animate DO and DO-cleft with inanimate DO. A disambiguating word was used in each item in order to clarify the nature of the objects, such as grateful and heavy. The response times of the disambiguating word as well as the six words following were measured.

It was found that readers spent significantly less time reading when forced to adopt a DO-cleft reading, but that within IO-clefts, participants spent significantly less time on those stimuli with an inanimate DO.

ExLing 2023 Athens: Proceedings of 14th International Conference of Experimental Linguistics, 18-20 October 2023, Athens, Greece

Conclusions and Implicatures Both predictions were borne out. The fact that IO is more difficult to displace is interesting in light of theories on distance effects (Gibson, 1998; Hawkins, 2004) since IO is usually both linearly closer to the front of the clause, and structurally higher and thus closer to its position after displacement. Moreover, strong evidence against modular theories (eg. Frazier, 1979) was also found in this research since semantic information, i.e., animacy, was clearly used by the parsers in our data to resolve the structural ambiguity.

References

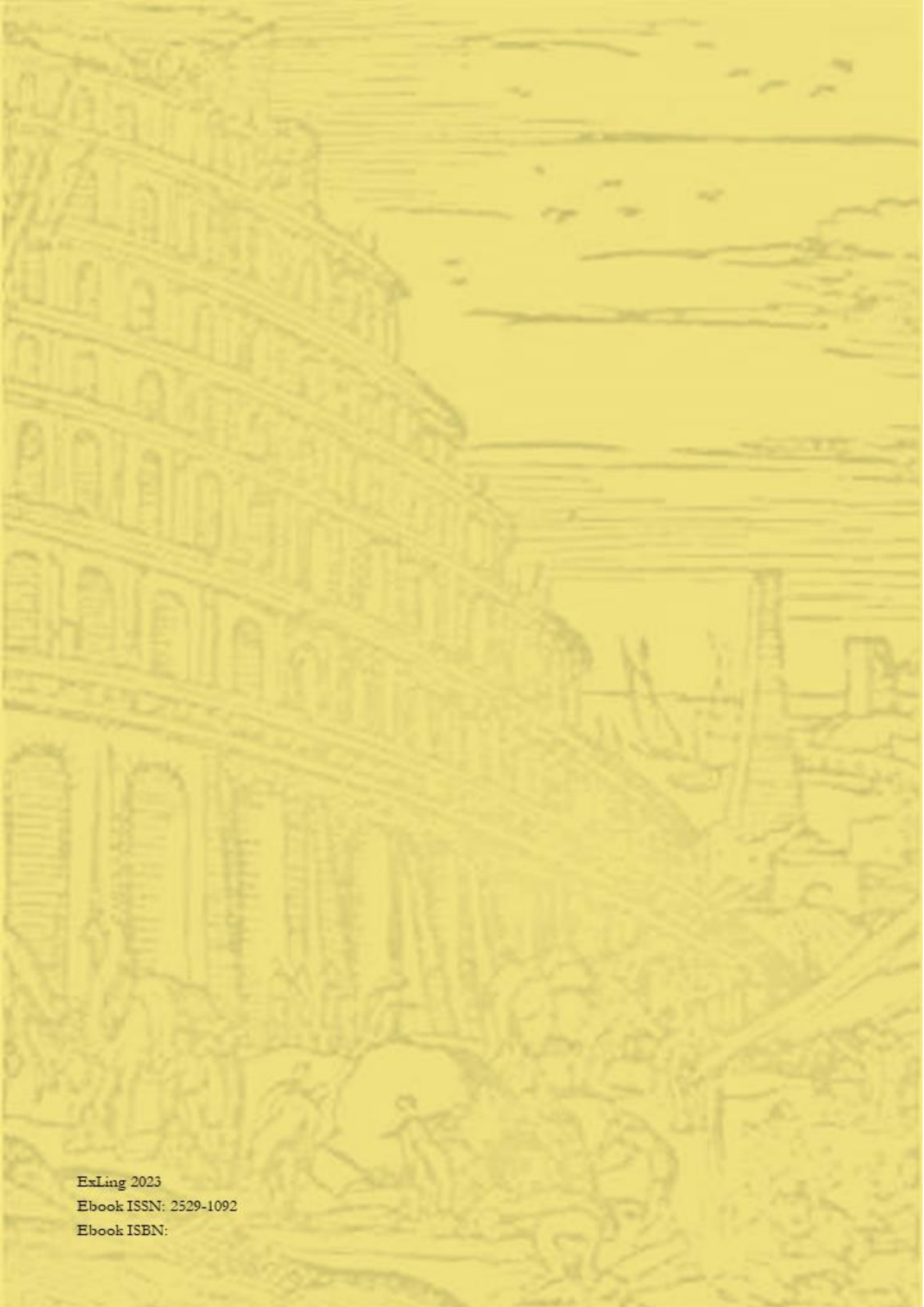
Frazier, L. 1979. On Comprehending Sentences: Syntactic Parsing Strategies. Ph.D. dissertation, Indiana University Linguistics Club, University of Connecticut. https://opencommons.uconn.edu/dissertations/AAI7914150.

Gibson, E. 1998. Linguistic complexity: locality of syntactic dependencies. Cognition, 68(1), 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1.

Hawkins, J.A. 2004. Efficiency and Complexity in Grammars. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199252695.001.0001.

Keenan, E.L., Comrie, B. 1977. Noun phrase accessibility and universal grammar. Linguistic inquiry, 8(1), 63-99. https://www.jstor.org/stable/4177973.

Index of names



ExLing 2023
Ebook ISSN: 2529-1092
Ebook ISBN: