



Prosodic disambiguation and attachment height

Nino Grillo¹, Giuseppina Turco²

¹Institut für Anglistik und Amerikanistik, Humboldt-Universität zu Berlin, Germany

²Laboratoire de Phonétique et de Phonologie (UMR 7018) CNRS/Sorbonne Nouvelle, France

nino.grillo@hu-berlin.de, giuseppina.turco@univ-paris3.fr

Abstract

This study investigates the role played by prosody in the syntactic and semantic disambiguation of string identical Relative Clauses (RC) and Pseudo Relatives (PR) in Italian. While RCs are embedded within the Noun Phrase (NP) they modify, PRs sit in a higher position in the syntactic structure, standing in a sisterhood relation with the head Noun. A production study with 8 Italian speakers set to determine whether and how this structural difference is encoded at a prosodic level. Preliminary results suggest that the minimal pairs are disambiguated at a prosodic level and that such disambiguation is encoded as early as at the NP-head. We discuss how this prosodic evidence reflects the structural difference highlighted above.

Index Terms: Prosodic disambiguation, attachment height, Italian (Pseudo) Relative Clauses

1. Introduction

The prosodic properties of various types of Relative Clauses (RCs) have been the object of detailed investigation [see 1, and references cited therein]. No work to date, however, has investigated the prosody of so called Pseudo Relative (PRs).

PRs (1) are finite embedded clause, available in Italian among other languages, which look superficially like RCs (2) but are naturally translated as English gerundive constructions (as illustrated in the translation to (1)). Like gerundives, they are licensed by a limited set of predicates, e.g. under perceptuals (1) they denote the direct perception of an event. The ambiguity disappears under stative verbs (3), which take NPs as complement and only allow a RC interpretation of the embedded *che*-clause (*che* meaning “that”; see [2, 3] for an overview on PRs).

- (1) Lia sentì [_{PR} [_{DP} il malato] [_{CP} che muoveva la barella]].
‘L. heard the patient moving the stretcher.’ PR
- (2) Lia sentì [_{DP} il [_{NP} malato] [_{CP} che muoveva la barella]].
‘L. heard the patient who moved the stretcher.’ RC
- (3) Lia curò il malato che muoveva la barella.
‘L. treated the patient who moved the stretcher.’ RC/*PR

A crucial distinction between PRs and RCs lies in the relation between the matrix verb and the ‘head-NP’ (*patient*). The matrix verb takes the NP as complement in (2) and (3). The complement of the matrix verb, however, is *not* the NP, but a clause (the PR) which contains the NP as subject and the Complementizer Phrase (CP) as predicate (1). As further shown in (1) and (2), while RCs are embedded within the NP they modify, the CPs of PRs stand in a sisterhood relation with the same NP.

These structural differences map into a number of interpretive differences, the most obvious being that, contrary to RCs,

the CPs of PRs are not integrated within the ‘head NP’ (*i.e. paziente*): they do not provide any information for its identification [4] and do not modify it in any way, not even by providing additional information about its referent, as Appositives do [1, 5]. As a result, RCs denote properties of individuals, while PRs denote events [3].

Since [6], the impressionistic comparison between Appositive RCs and PRs is a staple of the PR literature: uncontroversially, contrary to appositives, PRs do not involve a ‘comma intonation’ (*i.e.* in acoustic terms, a strong prosodic boundary). Nevertheless, to our knowledge, no study to date has investigated the prosody of PRs and, more generally, prosody is completely absent from the discussion of the distinction between PRs and *Restrictive RCs*.

Our question is whether and, if so, how these syntactic and semantic differences are encoded at a prosodic level.

2. Background

In a production study, [7] tested whether Italian speakers disambiguated syntactically ambiguous utterances (among other cases) and which strategies they employed. First, it was found that the prosodic resolution of ambiguous attachment of prepositional phrases, adverbial and relative clauses was the least consistent, compared to semantic ambiguity cases (*i.e.* scope of negation and focus operators). Second, among those cases that were prosodically disambiguated, speakers mostly varied intonational phrasing between the two possible readings. For instance, in a sentence like *Ha provato a mettersi in contatto col suo oculista a Roma*, the PP *a Roma* was in one case set off in an intermediate phrase (thereby evoking the reading “the subject was in Rome when he tried to get in touch with his ophthalmologist”, non-local, VP-attachment) and be part of a larger intermediate phrase in the other case (“the ophthalmologist and not the subject was in Rome”, local, NP-attachment). A similar strategy was applied to adverbial and relative clauses. For instance, in *La ragazza che mi ha fregato il taxi era una testimone dell’omicidio* (“the girl who stole my taxi was witness of the homicide”) speakers produced a (L- or H-) phrase boundary between the noun and the relative clause when non-restrictive meaning was intended.

More recently, [1] have provided further support to the common claim that Appositives RCs (ARC), but not Restrictive RCs (RRC), are separated by prosodic boundaries. The difference across RC types is consistent when the two are disambiguated grammatically. Importantly, the difference in boundary strength (but not in intonation) between ARCs and RRCs disappears when the latter are extraposed, which supports the idea that the difference in boundary strength depends on attachment site. The higher attachment site of ARCs [8] and Extraposed RCs [1]

correlates with stronger boundaries. This interpretation is also compatible with the findings in [7], where phrases with higher attachment sites were more likely to be set off from their surroundings than phrases with lower attachment sites were.

A similar pattern was also observed for RC-attachment in complex NPs (see [9, 10] and much related literature). Recent work on the topic, however, shows a much bigger role of prosodic prominence (with respect to phrasing) and a fundamental role of prosodic typology (see discussion in [11] and reference cited therein). A thorough comparison with the complex literature on RC-attachment in complex NPs is far beyond the scope of this short paper. We will therefore focus predominantly on the comparison with the literature discussed above.

Given that the attachment site of the CP of PRs is higher than that of RCs, we might expect to find similar prosodic differences to correlate with attachment height in the disambiguation of PRs and RCs. However, the overlap between previous literature and the structure under investigation is only partial. Three important factors distinguish the present structures from those previously studied in the literature. First of all, the verb-complement relation holding between the matrix predicate and the embedded NP in RCs of all kinds (e.g., restrictive, extraposed or not, and appositive) does not hold in PRs. Rather, the embedded NP of PRs stands in a subject-predicate relation with the embedded CP. Second, while it is true that attachment height varies across PRs and in a way which is reminiscent of previously studied cases (i.e. the CP sits in a structurally higher position than the one it occupies in run-of-the-mill non-extraposed restrictives), the similarity is not complete as only in PRs the NP and CP stand in a very local relation (sisterhood, i.e., the *most local* relation possible). On the other hand, all previous studies on the prosodic effect of attachment site involve a non-local *structural* relation between the NP and the CP (appositives are only linearly adjacent to the NP they modify, structurally their relation is non-local). On the relevance of locality for prosody see [1]. Last but not least, the attachment site of PRs is also *lower* than that of both Appositives and Extraposed RCs and the PR reading needs to be disambiguated from *both* the restrictive and the appositive/extraposed RC reading. This brings to the prediction that the height of attachment of the CP in PRs will be prosodically encoded by means that might only partially overlap with those of appositives or extraposed RCs.

3. The study

3.1. Methods

In a production study we compared the prosodic properties of utterances evoking either a PR or a RC reading (cf. examples above (1)/(2), “PR/RC ambiguous” condition henceforth) with phonetically similar utterances evoking a RC reading only (“RC unambiguous” condition (3)). These sentences were embedded in short paragraphs. The context introducing these sentences was designed to ensure a PR/RC reading and an unambiguous RC reading respectively, along with the manipulation of the matrix verb (cf. below).

3.2. Materials

The material comprised 24 experimental utterances (12 per condition) structured as follows: Noun Phrase (NP) matrix subject + matrix-V(erb) + NP matrix object + the CP (*who*-embedded V

+ NP object). These minimal pairs differed in the properties of the matrix verb only, while the region of interest (i.e. NP matrix object + CP) was kept identical. More specifically, utterances elicited in RC unambiguous condition contained stative predicates (e.g. *live with*) that can only take NPs as complements (3). This ensured that the following CP could be unambiguously interpreted as an RC. Moreover, to satisfy the contextual requirements of RCs, two alternative referents were introduced and a set of distinguishing features were specified for both of them (4).

(4) RC-CONTEXT:

During the weekend, we only had two patients in the clinic, one of them was calm, the other one didn't stop fidgeting for a second. **Target sentence (3).**

By contrast, PR/RC-compatible structures contained perceptual (e.g., *see, hear*) or semi-perceptual (e.g., *film, record*) verbs. Yet, previous comprehension studies have shown a strong parsing preference for the PR-interpretation in the environment of perceptual verbs [12, 13, 14, 15]. Thus, while a PR/RC ambiguity still exists under perceptual verbs, we can rely on previous results to the effect that in this environment the ambiguity will be resolved in favour of the PR interpretation. To match the RC unambiguous condition, and avoid *previous-mention* confound, the ‘head-NP’ (e.g. *patient*) was introduced also in the context of the PR/RC condition. In contrast to the RC unambiguous condition, the context introducing PR/RC ambiguous utterances contained only one referent (5).

(5) PR-CONTEXT:

We have a patient who's constantly moving. At the hospital we have clear rules as to where to keep everything. This is especially for safety reasons and to avoid hindering in case of emergency. **Target sentence (1).**

The minimal pairs were also controlled prosodically across items. The number of syllables did not differ of more than three syllables across items (between 16 and 19 syllables). All of the NP matrix objects, the embedded-Vs and the embedded-NP objects were trisyllabic, with the lexical stress on the penultimate syllable. Finally, creakiness or pitch lowering was avoided by following each experimental sentence by one or two additional sentences. The 24 experimental items were interspersed with 48 fillers. Fillers matched items in length and contained 4 sets of 12 sentences with different syntactic structures (i.e. Actives, Passives, Clefts and Clitic Left Dislocation).

3.3. Participants

Eight Italian native speakers originating from different regions participated in the experiment (age range= 19-to-48, age average=26.5, SD=8). Participants gave their informed consent and were given a small fee for their participation.

3.4. Procedure

Participants were instructed to read the paragraphs fluently, at normal speed, and silently scan the entire paragraph before reading aloud. Experimental stimuli were divided in two lists by using a Latin square design, such that each subject only produced one version of each sentence. This measure was taken to avoid repetition effects (e.g., deaccentuation of repeated words within items) that might introduce confounding. Stimulus sen-

tences and fillers were pseudo-randomized, automatically presented on a computer screen and recorded on a PC run using the Prompt and Record software ProRec 2.2 (©Mark Huckvale, University College London). The materials were recorded in a soundproof booth at experimental facilities at University College London. Each subject underwent a session comprising 72 items that lasted approximately 35 minutes.

3.5. Data analysis

Preliminary annotations were performed manually and blinded to the condition the sentence belonged to. Utterances were segmented and annotated on the syntactic, word and syllable level using Praat [16]. We measured some of the acoustic (temporal and melodic) properties of the target utterances, by largely following previous studies [1]. These were:

- the total duration (in ms) of the NP matrix object plus the CP; and separate temporal measures for the NP matrix object, the raw duration of the CP (calculated from the modal voicing offset of the previous vowel, i.e. the last vowel of the NP matrix object). Since log durations yielded the same statistical results as the raw durations, in the following section we will report results relative to the raw durations only;
- the presence of a tonal movement produced on the NP matrix object;
- the pitch excursion (calculated in st/sec) of the tonal movement produced on the NP matrix object;
- the presence of a tonal movement produced on the embedded-NP object;
- the pitch excursion (calculated in st/sec) of the tonal movement produced on the embedded-NP object.

Statistics were performed by using mixed effect regression models as implemented in the R-package lme4 [17]. For each dependant variable, we ran a model with CONDITION as fixed factor (PR/RC ambiguous vs. RC unambiguous), SPEAKER and ITEM as crossed-random factors, allowing for random intercepts and slopes by-participants and by-items adjustments for CONDITION [18, 19]. To control for potential artifacts, the models included phonemic length as an additional factor [1]. P-values were calculated on the basis of *Satterthwaite* approximation by using the lmerTest package [20].

3.6. Results

Of the 192 sentences (24x8=192), twelve were excluded because of hesitations or slight mispronunciations during the production of the target sentence (accounting for 6.2% of the data), thereby leaving 180 utterances for the analysis (91 utterances in the PR/RC ambiguous condition vs. 89 in the RC unambiguous condition). The values reported in Table 1 show that each target constituent is shorter in RC unambiguous than in PR/RC ambiguous condition. It furthermore suggests that the prosodic disambiguation already starts at the NP matrix object. The statistical model revealed a significant effect of CONDITION on the (combined) duration of the NP matrix object + the following CP ($\beta=-35.1$, $SE=12.9$, $t=-2.70$, $p=0.018^*$, the beta coefficient indicates the magnitude of the effect, here in milliseconds). Figure 1 shows the temporal differences of the whole region of interest (NP matrix object + CP) split by condition.

Table 1: Mean values and standard deviations of the duration (in ms) of the NP matrix object and of the CP in PR/RC ambiguous condition vs. RC unambiguous condition

	PR/RC ambiguous		RC unambiguous	
	Mean	SD	Mean	SD
NP duration	460.6	66.5	438.3	66.3
CP duration	1123.7	122.4	1092.0	102.7

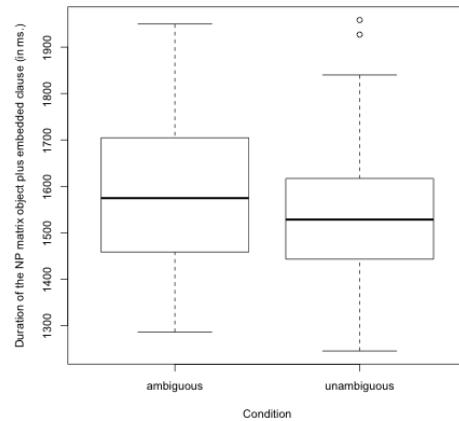


Figure 1: Average raw duration of the NP matrix object plus the CP (in ms) in (PR/RC) ambiguous condition vs. (RC) unambiguous condition.

As for the melodic properties of the NP matrix object, we observed that there were fewer (rising and falling) tonal movements in the RC unambiguous condition than in the PR/RC ambiguous condition. This distribution difference is illustrated in Figure 2. A binomial logistic regression analysis confirmed the main effect of CONDITION on accent distribution ($\beta=-2.73$, $SE=1.37$, $z=-1.98$, $p=.047^*$, 97.4% vs. 71.1%).

The pitch excursion (in st/sec) of the tonal movements illustrated in Figure 2 (the *yes*-coded cases only) was on average half of a tone narrower in RC unambiguous condition than in the PR/RC ambiguous condition. The model revealed that the excursion difference across conditions was only marginally significant ($\beta=-0.55$, $SE=0.28$, $t=-1.90$, $p=.07$).

Embedded-NP objects were typically produced with falling movements (73.8% of which were coded as H+L* -L%, in ToBI-style annotation). The analysis on the embedded-NPs shows no difference ($p=.2$) in accent distribution. In RC unambiguous condition, embedded-NP objects were accented in 82% of the cases ($n=65$); in PR/RC ambiguous condition, these were accented in 73% of the cases ($n=75$).

Finally, the acoustic analysis of the tonal movements (the falling movements only) showed no difference in pitch excursion, as also supported by the statistical model ($p=.8$).

4. Discussion and Future directions

The preliminary analyses of the production study suggest that on the whole Italian speakers make use of certain temporal

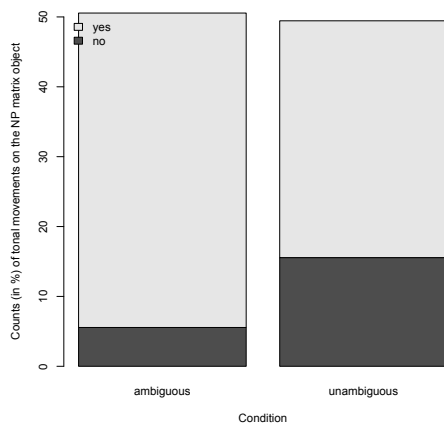


Figure 2: Presence (“yes”) and absence (“no”) of tonal movement on the NP matrix object in (PR/RC) ambiguous vs. (RC) unambiguous condition.

and melodic cues to disambiguate between PR and RC reading. More specifically, the longer duration of the NP matrix object may signal the presence of a major prosodic boundary. The temporal differences of the combined measure (i.e. duration of the NP matrix object plus the CP) may be taken to reflect the different structural relation between the matrix verb and the head-NP and between the head-NP and the CP in the sentences under consideration. The NP is the complement of the matrix verb in RCs, but the subject of the embedded clause in the PR. Similarly, the CP is embedded within the NP it modifies in the case of RCs, but sits in a higher attachment site in the case of PRs, where it stands in a sisterhood relation with the NP, in the case of PRs.

To the extent that these duration differences can be interpreted in terms of intonational phrasing differences, the current preliminary results are in line with previous work showing a similar interaction between prosody and syntax [7, 1, 21]. The present results add to this literature in two ways. First, they provide evidence to the general idea that higher attachment site correlates with separate phrasing (as evidenced by duration in our case); second, they hinge on a structural configuration previously untested: one in which the higher attachment site corresponds with a *local* structural relation between the head-NP and the CP. Previously tested structures all involved a non-local structural relation between the two elements in the higher attachment configuration. As discussed in [1], locality, or lack thereof, plays a major role in accounting for the prosodic properties of extraposition and appositives. While more work is needed to fully ascertain the similarities/differences between the PR/RC prosodic disambiguation and the more common cases studied in the literature (see below), our findings show that it might be possible to separate the contribution of different factors (i.e. attachment height and locality) in the prosodic realization of different types of embedded clauses.

Since our design included a comparison between unambiguous RC and ambiguous PR/RC condition, an alternative analysis based on predictability-based accounts of sentence production (see [22] among others) is also possible. Under this account, the observed duration differences between the two conditions would more simply indicate a difference in ambiguity (as sug-

gested by the data distribution in Figure 1). Specifically, the shorter duration observed in the RC unambiguous condition, as compared to the ambiguous PR/RC condition, would be attributed to its lack of ambiguity and, in a way, to a higher degree of predictability. However, as pointed out in the *Background* section, previous studies on the PR/RC ambiguity consistently found a strong preference for the PR-interpretation, which might reduce the strength of this alternative approach. Additional research with unambiguous PRs is thus necessary to adjudicate between the predictability and the attachment-site explanation.

It should be noted that nine of the utterances contained a silence at the onset of the complementizer *che* whose silence/closure duration was longer (average=126.9 ms, SD=89.4) than that of all the other cases (n=171, average=31.5 ms, SD=15.6). These nine utterances were judged to evoke a PR-adjunct reading by both authors, whereas an Appositive reading was excluded because of the lack of a (typical) separate intonational tune (see [1] for discussion). Interestingly, these cases occurred more frequently in PR/RC ambiguous (n=8) than in RC unambiguous condition (n=1). Besides the PR-complement interpretation ((1), in which the NP and the CP form a single constituent), a PR-adjunct parse also exists in which the embedded CP functions as a temporal modifier of the matrix VP (i.e. *while/at the time that*). PR-adjunct CPs attach even higher than those of PR-complements, i.e. as VP-modifiers (6) [2].

- (6) Lia [_{VP}[_{VP}beccò il malato][_{PR}che muoveva la barella]].
‘Lia caught the patient moving the stretcher.’

We speculate that this effect might have been induced by the specific pragmatic properties of certain items. For example, sentence 4 in the PR/RC condition reads: *Mario ascoltava la formaia che informava i cornetti*, ‘M. listened to the baker that baked the croissants’. The PR-complement interpretation would derive the slightly odd meaning in which Mario is listening to the event of baking. In the alternative PR-adjunct interpretation, the complement of the matrix verb is not the baking event, but the baker herself. The meaning of the whole clause would then be the more natural: Mario listened to the baker, while she baked the croissants. What is interesting in this result, which admittedly lacks quantitative significance, is that the higher attachment site of PR-adjunct with respect to PR-complement seems to correlate with an even more marked temporal difference.

In light of these observations, we are conducting further studies based on grammatically-driven disambiguation between the three readings (RC, PR-complement and PR-adjunct). Grammatical constraints might indeed be more effective than contextual disambiguation also for the cases tested here, in line with [1]’s observations. Along with that, more in-depth analyses of the current data looking at inter-speaker prosodic variation is also being carried out.

One difficulty with grammatically-driven disambiguation between PR and RC interpretation is its possible repercussions on prosody for independent reasons. For instance, one typical way to force the PR-parse is by means of pseudo-clefting: (*what I saw is the boy running*), which is likely to induce independent prosodic effects. Grammatical disambiguation, therefore, was discarded in this preliminary study which aimed at investigating the prosody of PRs vs. RCs in a more spontaneous experimental setting. Recently, we have found a way to by-pass this problem by forcing the PR-interpretation as in (7).

- (7) L’evento più strano della festa è l’uomo che balla. *RC
‘The strangest event of the party is the man dancing.’

5. Acknowledgements

This research was funded by the DFG Leibniz PrizeAL554/8-1 to Artemis Alexiadou (Nino Grillo) and by the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 662530 awarded to the second author.

References

- [1] C. Poschmann and M. Wagner, "Relative clause extraposition and prosody in german," *Natural Language Linguistic Theory*, pp. 1–46, 2015. [Online]. Available: <http://dx.doi.org/10.1007/s11049-015-9314-8>
- [2] G. Cinque, "The Pseudo-Relative and Acc-ing constructions after verbs of perception," in *University of Venice Working Papers in Linguistics*. Università di Venezia, 1992.
- [3] K. Moulton and N. Grillo, "Pseudo-relatives: Big but Transparent," in *NELS 45 (45th annual meeting of the North East Linguistics Society)*, vol. 2, MIT Cambridge, MA, 2014, pp. 193–202.
- [4] I. Heim and A. Kratzer, *Semantics in Generative Grammar*. Oxford: Blackwell, 1998.
- [5] A. Riester, "Stress test for relative clauses," in *Focus at the Syntax Semantics-Interface*, ser. SinSpeC. Working Papers of the SFB 732, E. Onea and A. Riester, Eds. Universitt Stuttgart, 2009, vol. 3, pp. 69–86.
- [6] A. Radford, *Italian Syntax. Transformational and Relational Grammar*. Cambridge: Cambridge University Press, 1977.
- [7] J. Hirschberg and C. Avesani, "The role of prosody in disambiguating potentially ambiguous utterances in english and italian," in *Intonation: Theory, Models and Applications*, 1997.
- [8] J. D. McCawley, "The syntax and semantics of english relative clauses," *Lingua*, vol. 53, no. 2, pp. 99–149, 1981.
- [9] J. D. Fodor, "Psycholinguistics cannot escape prosody," in *Speech Prosody 2002, International Conference, 2002*.
- [10] S.-A. Jun, "Prosodic phrasing and attachment preferences," *Journal of psycholinguistic research*, vol. 32, no. 2, pp. 219–249, 2003.
- [11] S.-A. Jun and J. Bishop, "Prominence in relative clause attachment: Evidence from prosodic priming," in *Explicit and Implicit Prosody in Sentence Processing*, ser. Studies in Theoretical Psycholinguistics, L. Frazier and E. Gibson, Eds. Springer International Publishing, 2015, vol. 46, pp. 217–240. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-12961-7_12
- [12] N. Grillo and J. Costa, "A novel argument for the universality of parsing principles," *Cognition*, vol. 133, no. 1, pp. 156–187, 2014.
- [13] N. Grillo, J. Costa, B. Fernandes, and A. Santi, "Highs and Lows in English Attachment," *Cognition*, vol. 144, pp. 116–122, 2015.
- [14] N. Grillo, B. Hemforth, C. Pozniak, and A. Santi, "Pseudo Relatives are easier than Relative Clauses: Evidence from Tense," in *28th CUNY Conference on Human Sentence Processing*, E. Kaiser, T. Mintz, R. Pancheva, and J. Zevin, Eds. Los Angeles: University of Southern California, March 2015.
- [15] —, "Pseudo Relatives are easier than Relative Clauses: Eye Tracking evidence from Tense," in *AMLaP (Architectures and Mechanisms for Language Processing)*, 2015.
- [16] P. Boersma and D. Weenink, "{P}raat: doing phonetics by computer (version 5.3.34)," 1992–2012, retrieved from <http://www.praat.org>.
- [17] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *arXiv preprint arXiv:1406.5823*, 2014.
- [18] D. J. Barr, R. Levy, C. Scheepers, and H. J. Tily, "Random effects structure for confirmatory hypothesis testing: Keep it maximal," *Journal of memory and language*, vol. 68, no. 3, pp. 255–278, 2013.
- [19] I. Cunnings, "An overview of mixed-effects statistical models for second language researchers," *Second Language Research*, vol. 28, no. 3, pp. 369–382, 2012.
- [20] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package)," *R package version*, vol. 2, no. 6, 2013, retrieved 15-11-2015, from <https://cran.r-project.org/web/packages/lmerTest/>.
- [21] M. Wagner, "Prosody and recursion in coordinate structures and beyond," *Natural Language & Linguistic Theory*, vol. 28, no. 1, pp. 183–237, 2010.
- [22] T. F. Jaeger, "Redundancy and reduction: Speakers manage syntactic information density," *Cognitive psychology*, vol. 61, no. 1, pp. 23–62, 2010.