Prosody of classic garden path sentences:
The horse raced faster when embedded

Nino Grillo[1], Miriam Aguilar[2], Leah Roberts[3] Andrea Santi[4], Giuseppina Turco[5]

[1]Department of Language and Linguistic Science, University of York
[2] Centro de Linguística da Universidade Nova de Lisboa
[3] Department of Education, University of York
[4]Department of Linguistics, University College London
[5]Laboratoire de Linguistique formelle (LLF) CNRS/Université Paris 7 Paris Diderot, France

`nino.grillo@york.ac.uk`, `maguilar@fcsh.unl.pt`, `leah.roberts@york.ac.uk`,
`a.santi@ucl.ac.uk`, `giuseppina.turco@univ-paris-diderot.fr`

Abstract

Prosody, it is assumed, does not always disambiguate syntax. We investigate one classic case at point from the psycholinguistics literature: garden path sentences involving the main-verb vs. reduced relative clause contrast (*the horse raced past the barn (and) fell*). Despite their centrality in shaping theories of sentence processing, no experimental work to date has investigated the prosody of these sentences. We show that, contrary to previous assumptions [1, 2], this contrast *is* prosodically disambiguated, but that this disambiguation can only be observed when the relevant clauses are embedded within a matrix clause which provides a baseline pace. Prosodic disambiguation obtains through pace modulation, with *faster* pace associated with the embedded/reduced relative reading and *regular* pace (no change) with main verb analysis. The essential contribution of the matrix sentence is to provide a baseline pace without which it is impossible to establish whether a change took place. Importantly, duration is solely determined by prosody and independent from complexity: faster pace is associated with the more complex structure.

Index Terms: Prosodic disambiguation, pace, complexity, garden-path sentences, embedding vs. sisterhood

1. Introduction

In current psychological models, and our everyday intuition, a simple correlation exists between relative task complexity and completion duration (when successful). Since Donders experiments in 1867, (reaction/response) time measures have been correlated with complexity and, in combination with other behavioural measures (i.e., accuracy), have consistently provided key insights into processes and mechanisms of the mind. We argue that, while generally sound, in the domain of language, and in particular when prosodic effects on duration are taken into account, this simple correlation can lead to dangerous over-simplifications. Recent psycholinguistics research shows that (Explicit and Implicit) prosodic properties, including phrasing, accentuation and rhythm, play a central role in sentence processing [3]. This work shows that prosody modulate durational properties of words and phrases to reflect their structural and interpretive properties. We claim that these effects can lead to apparently paradoxical cases of shorter durations for more complex structures. Prosody, it is assumed, does not always disambiguate syntax, the contrast between Main verb and Reduced-Relative Clause (RC) analysis in (1-a,b), is one classic case of such mapping failure:

(1)   a.   The horse raced past the barn and fell.
      b.   The horse raced past the barn fell.

We present evidence from a production study, that the classic garden path sentences in (1) are prosodically distinct and more generally that there exists a well-defined environmental contrast in which higher complexity co-occurs with shorter production/reading durations: a verb-phrase in a sisterhood vs embedded relation to a DP.

2. Background

Since the seminal work of Bever 1970, the contrast in (1), a case of local ambiguity between the Main Verb/Reduced RC parse, constitutes possibly the most well-known and one of the best studied examples of syntactic ambiguity in the literature and has provided one of the main testing grounds for different theories of sentence processing and of the relative contribution and timing of its subcomponents. The higher complexity of the reduced-RC analysis is not under discussion, rather the debate has focused on the underlying causes of this complexity and their relevance for sentence processing models. A variety of factors (including lexical, semantic, pragmatic and contextual) has been shown to modulate the strength of the garden path effect [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15].

Since the seminal work of Bever 1970, Despite this wealth of data, no study, to our knowledge, has investigated experimentally whether and how the reduced-RC vs. Main Verb ambiguity is prosodically disambiguated, although the general assumption has been that the relevant structural and interpretive differences are not prosodically encoded [16, 2]. While there are no a priori reasons to assume that prosody *always* disambiguates syntax, lack of prosodic disambiguation between the two readings is somewhat surprising, as previous results on similar structural contrasts from different languages clearly show a correlation between attachment height and intonational phrasing (see [17, 18, 19] a.o.).

The crucial distinction between Main Verb and the reduced RC in (1) lies in the relation between the DP (*the horse*) and VP (*raced past the barn*). The verb (the whole VP in fact) is embedded within the DP it modifiers in (figure 1[B]) but stands in a sisterhood relation with the same DP in (figure 1[A]).

Similar differences in attachment height have previously been shown to be prosodically encoded, with higher attachment site correlating with separate phrasing. [18], for example, provided further support to the common claim that Appositives
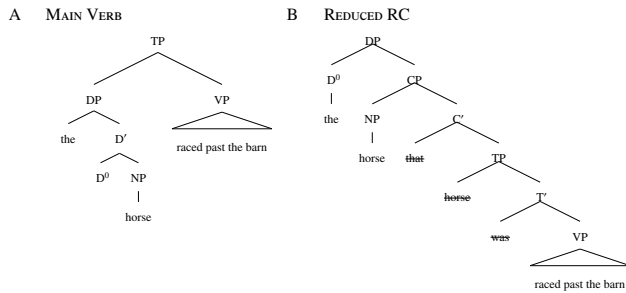
Figure 1: *Main Verb vs. Reduced RC parse. The present argument does not hinge on the specific RC analysis we adopted.*

Figure 2: *Pseudo Relative vs. Relative Clause parse.*

RCs, but not Restrictive RCs, 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 appositive and restrictive RCs 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 appositive RCs [20] and Extraposed RCs correlates with stronger boundaries. This interpretation is also compatible with the findings in [17], where phrases with higher attachment sites were more likely to be set off from their surroundings than phrases with lower attachment sites were. [19] specifically tested the prosody of embedding (RCs) vs. sisterhood (Pseudo Relatives) and show that speakers make use of both temporal (shorter duration for embedded material) and melodic cues to encode this distinction. [19] investigated the role played by prosody in the disambiguation of string identical Relative Clauses (RC) and Pseudo Relatives (PR) in Italian. PRs (2) are finite embedded clause, available in Italian among other languages, which look superficially like RCs (3) but are naturally translated as English eventive Small Clauses (as illustrated in the translation to (2)). Like eventive Small Clauses, PRs are licensed by a limited set of predicates, e.g. under perceptuals (2) they denote the direct perception of an event, unlike RCs which are allowed under all sorts of predicates and denote properties of entities.

(2)   Ho visto il cavallo che correva.
        I.have seen the horse that was running.
        'I saw the horse running.'   PR

(3)   Ho visto il cavallo che correva.
        I.have seen   the   horse   that was running .
        'I saw the horse that was running.'   RC

The contrast is very similar to the one investigated here: while the CP of RCs is embedded within the Determiner Phrase they modify, the CP sit in a higher position in the syntactic structure in Pseudo Relatives, standing in a sisterhood relation with the head Noun (figure 2), in subject/predicate relation much like the one instantiated by the Main Verb analysis in figure 1[A]. Moreover, the Pseudo Relatives parse, just like the Main Verb parse, has been previously shown to be easier to parse than the RC parse [21, 22, 23, 24]. [19] show that PRs were associated with longer duration at both the NP and the CP, an observation which the authors link to the different structural position of the CP in the two readings, and take to reflect differences in intonational phrasing between the two readings, adding to a growing literature on the relation between duration and at-
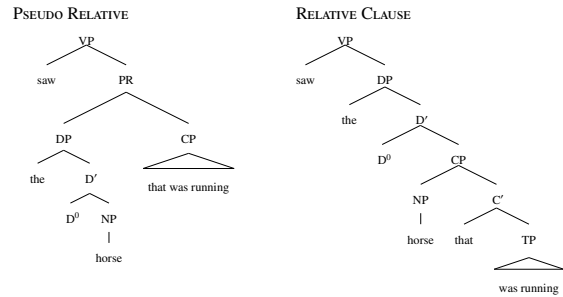
tachment height [17, 18, 25]. Higher attachment site correlates with separate phrasing, this is often observable in terms of durational differences between the two readings, with *shorter* durations for more deeply embedded strings and *longer* durations for high attachment of the same string.

Our first question, therefore, is whether and, if so, how the syntactic and semantic differences between the two readings of classic garden path sentences of the Main Verb/Reduced-RC variety are encoded at a prosodic level. More specifically, we ask whether the structural differences (i.e. differences in attachment height) between the position of the verb (raced) with respect to the subject DP (horse) results in prosodic differences.

We show that, contrary to previous assumptions [16, 2], the contrast in (1) *is* prosodically disambiguated, as early as the subject DP (*the horse*), aligning the Main Verb/Reduced RCs with other cases of attachment height disambiguation. This disambiguation, however, is best observed when the ambiguous string is embedded within a matrix clause which provides a baseline pace. Prosodic disambiguation obtains through pace modulation, with *faster* pace associated with the embedded/reduced relative reading and *regular* pace (no change) with main verb analysis. The essential contribution of the matrix sentence is to provide a baseline pace without which it is impossible to establish whether a change took place. Importantly, duration is solely determined by prosody and independent from complexity: faster pace is associated with the more complex structure.

3. The study

3.1. Methods

In a planned production study we compared the prosodic properties of utterances evoking a Main Verb reading with phonetically similar utterances evoking a reduced-RC reading only. These sentences were embedded in short introductory sentences containing declarative verbs. The short introduction was neutral with respect to the relevant disambiguation and was present solely to provide a baseline tempo. Notice that, for convenience, we still refer to "Main Verb" parse, even though in the materials used the relevant verb is part of the embedded declarative clause.

3.2. Materials

The material comprised 16 experimental utterances per condition adapted from previous experiments in the relevant garden path literature [7, 11]. Each experimental sentence was structured as follows: Noun Phrase (NP) matrix subject + matrix-

verb + that + ROI (DP + VP, in *italics* in (4)) + disambiguating coda (*(and) got badly hurt*).

(4)   MAIN VERB CONDITION:
      Jason claims that [$_{TP}$[$_{DP}$*the student*] [$_{VP}$*pushed into the row of traffic*] and [$_{VP}$*got badly hurt*]].

(5)   REDUCED-RC CONDITION:
      Jason claims that [$_{TP}$ [$_{DP}$*the* [$_{NP}$*student* [$_{CP}$*pushed into the row of traffic*] [$_{VP}$*got badly hurt*]]].

These minimal pairs differed in the properties of coda only, while matrix clause and the region of interest (ROI: i.e. embedded DP + VP, *the student pushed into the row of traffic*) were kept identical across conditions. More specifically, only utterances elicited in Main Verb condition contained a conjunction immediately after the ROI, which allowed to coordinate the second VP (*and got badly hurt*) with the first one and avoid the reduced-RC reading. The coda also provided a mean to avoid creakiness or pitch lowering in the ROI. The experimental items were interspersed with 48 fillers. Fillers matched items in length and contained sentences with different syntactic structures (i.e. Actives, Passives and Clefts).

3.3. Participants

Five English native speakers participated in the experiment (age range= 19-to-34, age average=25.3, SD=8). Participants gave their informed consent and were paid for their participation. Each subject participated in the experiment twice, with at least two weeks between each session.

3.4. Procedure

Participants were instructed to read the sentences fluently, at normal speed, and silently scan the entire sentence 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 in each session. This measure was taken to avoid repetition effects (e.g., deaccentuation of repeated words within items) that might introduce confounding. In a second session, at least two weeks away from the first one, each subject produced the alternative version of each sentence, which provide us with a total of 32 sentences per subject. Stimulus sentences 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 of York. Each subject underwent a session comprising 64 items that lasted approximately 35 minutes.

3.5. Data analysis

Segmentation was performed automatically by using SPPAS software [26]. F0 detection was automatically detected by means of scripts run in Praat software. The results of the automatic procedure was checked and manually corrected (blinded to the condition the sentence belonged to) in case of errors.
For each sentence, the following acoustic properties were measured:

- the total duration (in ms) of the (head) NP subject;
- the total duration (in ms) of the ROI. Since normalized duration yielded the same statistical results as the raw duration, in the following section we will report results relative to the raw duration only;

- building on previous studies (e.g. [27]) it is predicted a major difference in the reset of the pitch should be observed for the Main Verb condition compared to the reduced-RC condition after the ROI.

Statistics were performed by using mixed effect regression models as implemented in the R-package lme4 [28]. For each dependent variable (cf. above), we ran a model with CONDITION as fixed factor (Main Verb vs. reduced-RC), SPEAKER and ITEM as crossed-random factors, allowing for random intercepts and slopes by-participants and by-items adjustments for CONDITION [29, 30]. P-values were calculated on the basis of *Satterthwaite* approximation by using the lmerTest package [31].

3.6. Results

Of the 160 sentences (16 sentences x 2 conditions x 5 speakers), Five were excluded because of hesitations or slight mispronunciations during the production of the target sentence, thereby leaving 155 utterances for the analysis. The values reported in Table 1 show that each target constituent is shorter in reduced-RC than in Main Verb condition. The statistical models revealed a significant effect of CONDITION on the duration of the ROI ($\tilde{\chi}^2(1)$=5.40, p=0.02*): the ROI was significantly shorter in reduced-RC condition than in Main Verb condition (β=-85.4, *SE*=31.3, *t*=-2.729, *p*=0.0155*). The same result holds for the duration of the head NP subject ($\tilde{\chi}^2(1)$=4.92, p=0.02*): the head NP was significantly shorter in reduced-RC condition than in Main Verb condition (β=-33.1, *SE*=13.6, *t*=-2.425, *p*=0.0167*)

Figure 1 and 2 show the temporal differences of the whole region of interest (*the student pushed into the row of traffic*) and of the head of the subject NP (*the student*) respectively split by condition.

As for the pitch reset, there was almost one semitone difference across the two conditions (higher in Main Verb than in reduced-RC condition). This difference however was only marginally significant ($\tilde{\chi}^2(1)$=3.20, p=0.07).

Table 1: *Mean values and standard deviations of the duration (in ms) of the ROI and the Embedded (head of) Subject DP in Main Verb vs. Reduced-RC condition*

	Main Verb		Reduced-RC	
	Mean	SD	Mean	SD
ROI duration	2199	463.7	2122	442.9
DP duration	682.5	237.5	645.6	231.2

4. Discussion

We tested the hypothesis that structural embedding is associated with shorter duration (faster speech rate with respect to a baseline set by the matrix clause) at the prosodic level, while longer duration (regular pace) is associated with higher attachment site (sisterhood). This proposal, based on previous results from a comparable contrast in Italian [19], makes an important prediction: the more complex reduced RC (1)[b] should be associated with shorter duration than the easier Main Verb analysis (1)[a]. This is somewhat surprising under the reasonable assumption that a rational speaker should be expected to produce more complex structures more slowly. Nevertheless, as predicted, shorter duration was observed for the more complex
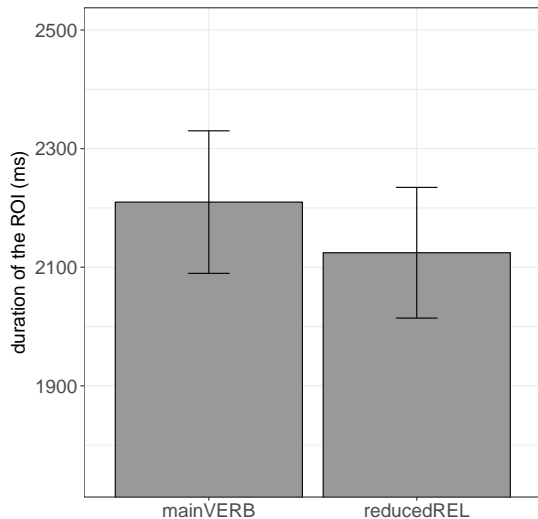
Figure 3: *Average raw duration of the ROI (in ms) in Main Verb vs. Reduced-RC condition.*

Figure 4: *Average raw duration of the (head of) subject DP (in ms) in Main Verb vs. Reduced-RC condition.*

reduced-RC condition, a result which comes with the important implication that listeners necessarily have less time to process the more complex reduced-RC. In other words, the domain of syntax-prosody interface presents an interesting case in which more complex structures have to be processed faster than simpler ones.

The preliminary analyses of the production study suggest that, contrary to previous assumptions [1, 2] on the whole English speakers make use of both temporal cues to disambiguate between Main Verb and reduced-RC readings and that this disambiguation is observable already at the head of the subject DP (*the horse*). To the best of our knowledge, previous claims that the two structures are not prosodically disambiguated are based on impressionistic judgments of the sentences in isolation. We have argued that indeed the temporal dimension of this disambiguation might go unnoticed when the sentences are presented in isolation and presented the relevant sentences as complement clauses of verba dicendi (*John said that ...*). The durational differences that constitute the main source of disambiguation of the two readings are not absolute, but relative to a baseline pace set by e.g. a matrix sentence (or alternatively by the preceding discourse). This is because shorter/longer duration of a string produced in isolation is not informative *per se*, as it might simply be taken to reflect aspect of speech rate irrelevant to syntax-prosody mapping. Just as previously observed with the Pseudo Relative/Relative Clause contrast, the longer duration of the head of the embedded subject in the Main Verb analysis may signal the presence of a major prosodic boundary. The temporal differences of the combined measures (duration of the DP plus the VP, i.e. the whole ROI) may be taken to reflect the different structural relation between the DP and the VP in the sentences under considerations. The DP and the VP form a single constituent in the reduced-RC analysis, with the VP (*raced past the barn*) embedded within, acting as a modifier of, the DP (*the horse*). Together, they form a modified subject of the embedded clause (*the horse fell*). In the Main Verb analysis, on the other hand, the DP (*the horse*) is the subject of an embedded clause which contains two conjoined VPs (*raced past the barn and fell*). In other words, while with reduced-RCs the relation between the VP and the DP is one of embedding, in the case of
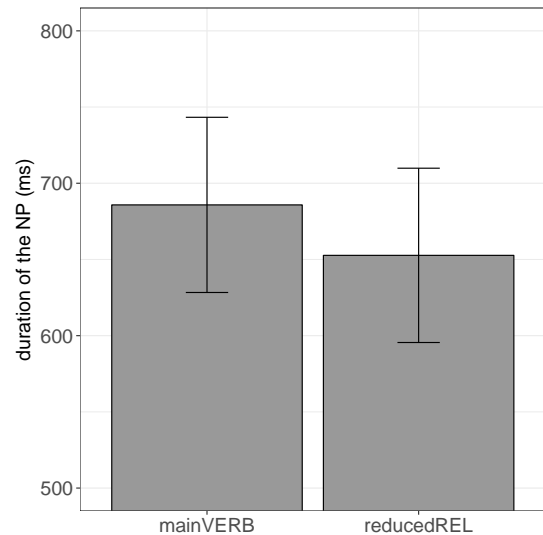
the Main Verb analysis it is one of sisterhood. To the extent that these duration differences can be interpreted in terms of intonational phrasing differences, the current preliminary results are in line with previous work showing a similar interaction between prosody and syntax [19, 17, 18, 25]. As mentioned in the introduction, in this study we chose to be faithful to original research on garden path effects by using sentences from previous studies which had been shown to derive strong complexity effects in the reduced-RC condition. This was important since we set to investigate the potential independence of duration from complexity. This choice, however, proved problematic for the analysis of tonal differences across conditions. The materials, while kept constant across conditions, varied greatly across items. The VPs of the embedded sentences varied in length (both in terms of number of words, syllables and characters), but also in terms of argument and event structure. We are currently working on follow up study with carefully controlled prosody across items.

5. Conclusions and Outlook

We tested the hypothesis that structural embedding is associated with shorter duration (faster speech rate with respect to a baseline set by the matrix clause) at the prosodic level, while longer duration (regular pace) is associated with higher attachment site (sisterhood). This proposal, based on previous results from a comparable contrast in Italian [19, 21], makes an important prediction: the more complex reduced RC (1)[b] should be associated with shorter duration than the easier Main Verb analysis (1)[a]. Preliminary results from a production study support this prediction. An important implication of these results is that listeners necessarily have less time to process the more complex reduced RC. In other words, the domain of syntax-prosody interface presents an interesting case in which more complex structures have to be processed faster than simpler ones. Future work is required to test the extent of this mapping across languages and structures and of its implications for psycholinguistics.

6. Acknowledgements

7. References

[1] J. D. Fodor, "Prosodic disambiguation in silent reading," in *PROCEEDINGS-NELS*, vol. 1, no. 32; VOL 1, 2002, pp. 113–132.

[2] M. Wagner and D. G. Watson, "Experimental and theoretical advances in prosody: A review," *Language and cognitive processes*, vol. 25, no. 7-9, pp. 905–945, 2010.

[3] L. Frazier and E. Gibson, *Explicit and Implicit Prosody in Sentence Processing: Studies in Honor of Janet Dean Fodor*. Springer, 2015, vol. 46.

[4] T. G. Bever, "The cognitive basis for linguistic structures," *Cognition and the development of language*, vol. 279, no. 362, pp. 1–61, 1970.

[5] J. Kimball, "Seven principles of surface structure parsing in natural language," *Cognition*, vol. 2, no. 1, pp. 15–47, 1973.

[6] L. Frazier, "Sentence processing: A tutorial review." 1987.

[7] S. Crain and M. Steedman, *On not being led up the garden path: the use of context by the psychological syntax processor*. United States: Cambridge University Press, 1985, pp. 320–358, cambridge Books Online.

[8] G. Altmann and M. Steedman, "Interaction with context during human sentence processing," *Cognition*, vol. 30, no. 3, pp. 191–238, 1988.

[9] F. Ferreira and C. Clifton, "The independence of syntactic processing," *Journal of memory and language*, vol. 25, no. 3, pp. 348–368, 1986.

[10] J. C. Trueswell, M. K. Tanenhaus, and S. M. Garnsey, "Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution," *Journal of memory and language*, vol. 33, no. 3, p. 285, 1994.

[11] W. Ni, S. Crain, and D. Shankweiler, "Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities," *Language and Cognitive Processes*, vol. 11, no. 3, pp. 283–334, 1996.

[12] K. B. Paterson, S. P. Liversedge IV, and G. Underwood, "The influence of focus operators on syntactic processing of short relative clause sentences," *The Quarterly Journal of Experimental Psychology: Section A*, vol. 52, no. 3, pp. 717–737, 1999.

[13] K. McRae, M. J. Spivey-Knowlton, and M. K. Tanenhaus, "Modeling the influence of thematic fit (and other constraints) in online sentence comprehension," *Journal of Memory and Language*, vol. 38, no. 3, pp. 283–312, 1998.

[14] C. Clifton, M. J. Traxler, M. T. Mohamed, R. S. Williams, R. K. Morris, and K. Rayner, "The use of thematic role information in parsing: Syntactic processing autonomy revisited," *Journal of Memory and Language*, vol. 49, no. 3, pp. 317–334, 2003.

[15] J. C. Sedivy, "Invoking discourse-based contrast sets and resolving syntactic ambiguities," *Journal of Memory and Language*, vol. 46, no. 2, pp. 341–370, 2002.

[16] J. D. Fodor, "Psycholinguistics cannot escape prosody," in *Speech Prosody 2002, International Conference*, 2002.

[17] 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.

[18] 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

[19] N. Grillo and G. Turco, "Prosodic disambiguation and attachment height," *Speech Prosody 2016*, pp. 1176–1180, 2016.

[20] J. D. McCawley, "The syntax and semantics of english relative clauses," *Lingua*, vol. 53, no. 2, pp. 99–149, 1981.

[21] N. Grillo and J. Costa, "A novel argument for the universality of parsing principles," *Cognition*, vol. 133, no. 1, pp. 156–187, 2014.

[22] N. Grillo, J. Costa, B. Fernandes, and A. Santi, "Highs and Lows in English Attachment," *Cognition*, vol. 144, pp. 116–122, 2015.

[23] 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.

[24] ——, "Pseudo Relatives are easier than Relative Clauses: Eye Tracking evidence from Tense," in *AMLaP (Architectures and Mechanisms for Language Processing)*, 2015.

[25] M. Wagner, "Prosody and recursion in coordinate structures and beyond," *Natural Language & Linguistic Theory*, vol. 28, no. 1, pp. 183–237, 2010.

[26] B. Bigi, "Sppas-multi-lingual approaches to the automatic annotation of speech," *"The Phonetician - International Society of Phonetic Sciences*, no. 111-112/2015, pp. 54–69, 2015.

[27] D. R. Ladd, "Declination "reset" and the hierarchical organization of utterances," *The Journal of the Acoustical Society of America*, vol. 84, no. 2, pp. 530–544, 1988.

[28] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *arXiv preprint arXiv:1406.5823*, 2014.

[29] 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.

[30] I. Cunnings, "An overview of mixed-effects statistical models for second language researchers," *Second Language Research*, vol. 28, no. 3, pp. 369–382, 2012.

[31] 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/.