

Interindividual variation in weighting prosodic and semantic cues during sentence comprehension – a partial replication of Van der Burght et al. (2021)

Constantijn L. van der Burght¹, Antje S. Meyer^{1,2}

¹Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

²Radboud University, Nijmegen, Netherlands

Stan.vanderBurght@mpi.nl, Antje.Meyer@mpi.nl

Abstract

Contrastive pitch accents can mark sentence elements occupying parallel roles. In “Mary kissed John, not Peter”, a pitch accent on *Mary* or *John* cues the implied syntactic role of *Peter*. Van der Burght, Friederici, Goucha, and Hartwigsen (2021) showed that listeners can build expectations concerning syntactic and semantic properties of upcoming words, derived from pitch accent information they heard previously. To further explore these expectations, we attempted a partial replication of the original German study in Dutch. In the experimental sentences “Yesterday, the police officer arrested the thief, not the inspector/murderer”, a pitch accent on subject or object cued the subject/object role of the ellipsis clause. Contrasting elements were additionally cued by the thematic role typicality of the nouns. Participants listened to sentences in which the ellipsis clause was omitted and selected the most plausible sentence-final noun (presented visually) via button press. Replicating the original study results, listeners based their sentence-final preference on the pitch accent information available in the sentence. However, as in the original study, individual differences between listeners were found, with some following prosodic information and others relying on a structural bias. The results complement the literature on ellipsis resolution and on interindividual variability in cue weighting.

Index Terms: pitch accents, individual differences, thematic roles, contrastive focus, sentence comprehension

1. Introduction

Pitch accents mark important or novel information in a sentence (for a review, see [1]). Contrastive pitch accents can be interpreted to contrast sentence elements. When asked, *Did you buy apples at the market?* the conversation partner may answer: *No, I bought BANANAS*. The capital letters here indicate the realisation of a contrastive pitch accent. In this way, the speaker corrects the information posited in the question, and presents contrastive information instead [2]. Listeners can use pitch accent information for reference resolution [3], [4], [5], [6], attachment ambiguity resolution [7], [8], or constituent order ambiguities [9]. As illustrated with the sentence *Mary kissed John, not Peter*, in ellipsis clauses contrastive pitch accents can be used to indicate elements occupying parallel syntactic roles [10], [11].

The resolution of ellipsis clauses is an interesting case, because, without verb, listeners need to derive its syntactic structure from other cues. In the example above, Peter can either be the implied subject or object of the ellipsis (i.e., he was not kissed by Mary vs. he did not kiss John). But which factors shape the interpretation that listeners arrive at? Both pitch accents, marking focus in the antecedent clause, and the

semantic parallels between noun phrases and are known to have effects [10], [12], [13], [14]. Additionally, listeners are known to have bias interpretations, which prosodic cues can only partly overcome [15].

A recent study investigated how listeners combine pitch accent information with syntactic and semantic cues to interpret ellipsis clauses [16]. In sentences such as “Yesterday, the policeman arrested the thief, not the murderer”, parallel roles were indicated by several cues available in the sentence: prosody (a contrastive pitch accent on either subject or object as well as the ellipsis noun phrase), semantics (thematic roles typicality), and syntax (in German, case marking of the determiner indicates subject or object roles). In the first experiment, sentences were constructed so that the various cues conflicted with one another, creating parallels that were syntactic or semantic mismatches (e.g., pitch accents establishing parallel roles for *policeman* and *murderer*, or *thief* and *detective*). Comprehension questions probed the subject or object interpretations of each noun phrase. The results indicated that case marking was the decisive cue, but that prosodic and semantic cues also influenced listener interpretations. In a second experiment, the same auditory sentences were interrupted after the main clause, after which listeners were asked to complete the ellipsis noun phrase. Results suggested that they based their choice of sentence ending on the pitch accents in the antecedent clause, selecting the noun phrase that was syntactically and semantically congruent to the focus-marked noun phrase. Together, the results suggested that pitch accents establish expectations concerning the semantic and syntactic features of the upcoming ellipsis clause.

In addition to these results on the group level, the study demonstrated considerable differences between comprehenders. Specifically, one group of listeners responded with good sensitivity to the pitch accent cues, whereas another group seemed to respond according to a structural bias. As in other language domains, there are known individual differences in prosody processing. In prosody production, variability between speakers in the realisation of question/statement intonation [17] and prosodic boundaries [18] have been reported. In prosody comprehension, interindividual variability has been shown in boundary perception [19]. In perception of prominence, it has been argued that listeners might differ in whether they pay more attention to pitch-related cues as compared to syntactic or semantic cues [20]. However, it is unclear which speaker and listener characteristics lead to greater sensitivity to prosody as compared to other cues.

Here, we report a replication attempt of Experiment 2 from Van der Burght et al. [16], translating the original German design into Dutch. We omitted the syntactic manipulation from the original study design and focused on the semantic condition

instead: while grammatical gender exists in Dutch, a case-marking manipulation similar to the one in the German study was not feasible as Dutch determiners are not marked for case. In the current study, participants listened to sentences of the type: *Yesterday, the policeman arrested the thief, not the murderer*. The subject/object role of the ellipsis clause was determined by thematic role typicality of the noun (a murderer is a likely patient of the verb *to arrest*). Additionally, a contrastive pitch accent on the subject or object of the antecedent clause (on *policeman* or *thief*) cued whether the constituent in the ellipsis would occupy a subject or object role. Crucially, the sentences were truncated before the ellipsis noun phrase (after *not*) and listeners were asked to complete the sentence: in a 2-alternative forced-choice task, they chose between two nouns that were either a typical agent or patient (presented visually) of the verb, thereby indicating whether they had interpreted a subject or object focus structure.

A replication study seemed worthwhile for several reasons. First, to assess whether the group-level results would replicate and generalise from German to Dutch. Considering that German and Dutch are largely similar in terms of pitch accent use, we hypothesised that listeners would indeed use pitch accent information to establish semantic properties of the ellipsis clause. The second aim was to explore whether we would find between-listener variability in responses, similar to the original study results. The origins of such differences were to be determined in further studies. As a final motivation, we considered the scientific value of replication studies in general and the relative scarceness of replications in psycholinguistics as compared to other fields [21].

2. Methods

The design mirrored Experiment 2 from the original paper [16] as closely as possible.

2.1. Participants

Participants (N=36) were native speakers of Dutch reporting no hearing or language disorders, recruited from the database of the MPI for Psycholinguistics. Most participants in the sample were part of the student population from Nijmegen and surrounding areas. All were paid 11€ for participation and gave informed consent prior to the experiment. The study was approved by the Ethics Committee Faculty of Social Sciences, Radboud University.

2.2. Experimental design

In our Dutch stimulus sentences, one out of two constituents in the main clause was contrastively focused (CF) with a third constituent in the ellipsis. The prosodic manipulation was established by a pitch accent (indicated with capital letters in the examples below), which marked whether focus was on the subject (1) or object (2) noun phrase of the antecedent clause. The semantic manipulation was established by the thematic role typicality of the nouns in relationship to the verb [22]: nouns were either typical agents (e.g., police officers) or patients (e.g., thieves) associated with the verb (e.g., to arrest). As a consequence, the subject/object role of the noun in the ellipsis was cued by prosodic as well as semantic properties. All nouns belonged to the common (shared masculine/feminine) gender category.

1. Gisteren heeft [de POLITIEAGENT]_{CF} de dief gearresteerd, niet [de INSPECTEUR]_{CF}

Yesterday the POLICE OFFICER arrested the thief, not the INSPECTOR

2. Gisteren heeft de politieagent [de DIEF]_{CF} gearresteerd, niet [de MOORDENAAR]_{CF}
Yesterday the police officer arrested the THIEF, not the MURDERER

We created 48 experimental sentences, which were recorded by a female native speaker of Dutch, who was instructed to realise a contrastive accent (corresponding to L+H* in the ToBI system) on either subject or object of the main clause, as well as the sentence-final noun phrase. The digitized speech signal (sampling rate 44.1 kHz; resolution 16 bits) was edited in Praat [23], removing the sentence-final noun phrases. The two sentence-final noun phrases (one agent-like and one patient-like noun) were presented in written form as response options.

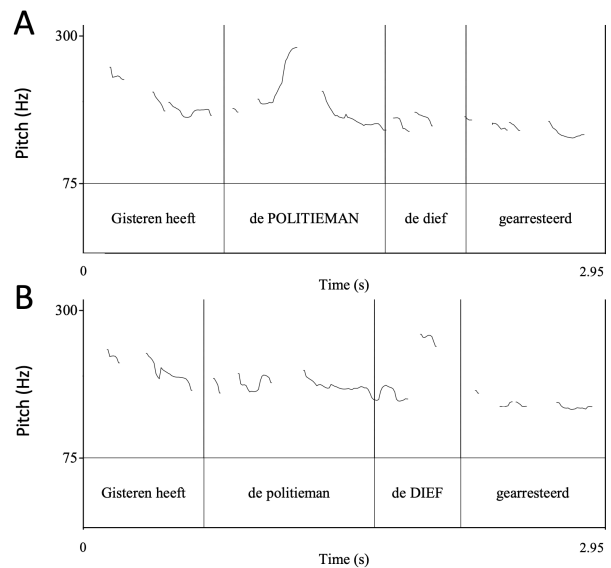


Figure 1: *Pitch contours of the subject (A) focus and object (B) focus conditions of the sentence “Yesterday the policeman arrested the thief”*

The materials included filler sentences (50% of all trials), which were spoken in neutral prosody (i.e., with regular accentuation, but without contrastive pitch accents). They were designed to mirror the experimental sentences in terms of task demands, but without the requirement of interpreting sentence prosody to perform the task. The fillers were short, transitive sentences of the structure SVO, ending in a direct object (e.g., “Het meisje aait het konijn”, *The girl pets the rabbit*) or prepositional clause (e.g., “De turner springt op de trampoline”, *The gymnast jumps on the trampoline*). Analogously to the experimental materials, fillers sentences were truncated after the verb and listeners indicated their preferred sentence ending based on semantic plausibility. Each sentence-final word was paired with a semantically implausible alternative (e.g., *pencil* or *conductor*, respectively, in the examples above).

The experimental items had been normed using an online survey to select items with the strongest thematic role hierarchy between nouns and verb. In this norming study, sentences were presented (visually) in both plausible and implausible form (i.e., *the police officer arrested the thief* vs. *the thief arrested the police officer*) [24]. Participants were asked to rate the sentences from 1 (completely implausible) to 7 (completely

plausible). There were two lists, so that the plausible and implausible versions of the same item were not normed by the same participant. From the 73 items normed, the 48 items with the largest difference between the plausible and implausible version were selected to be used in the experiment. Participants in the norming study were excluded from participation in the actual experiment.

2.3. Procedure

Participants sat in a sound-proof chamber in front of a computer screen and response button box. They started with a short practice session that was identical to the actual experiment but contained different stimuli. Each experimental trial (Figure 2) started with the presentation of a fixation cross. The auditory stimulus was played, followed by the visual presentation of the response options. Participants chose the noun phrase on the left or right side of the screen by pressing the left or right button. After the button-press response the next trial started.

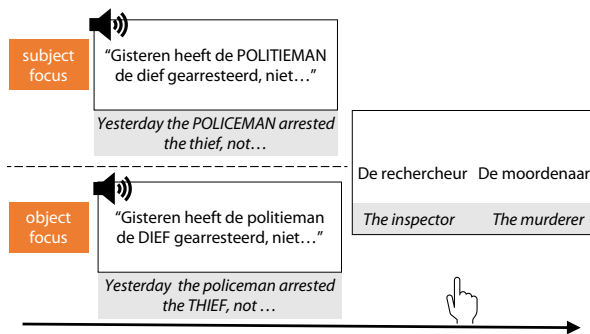


Figure 2: Schematic representation of the design. Participants were presented with a truncated sentence and were asked to complete the sentence with the most appropriate noun phrase.

Trials were presented in pseudo-randomised order with the constraint that trials of the same type (experimental, filler) could be presented no more than three times after another. The allocation of the agent and patient nouns to the left or right side of the response screen was counterbalanced across participants. The experiment was divided into four blocks, with self-timed breaks in between (minimum duration of 20 seconds). The total experiment duration was about 25 minutes. Trial presentation was controlled by Presentation software (Neurobehavioral Systems).

2.4. Statistical analysis

The proportions of agent/patient responses were analysed with a Generalised Linear Mixed Model (GLMM), using the LME4 package (version 1.1-30) in R (version 4.2.1). The full model included focus as fixed effect and by-participant as well as by-item random intercepts and slopes for focus.

Additionally, we performed a signal detection theory (SDT) analysis [25] to dissociate sensitivity to the prosodic cues (d') and response bias. We treated the subject-focus trials as 'signal' and object-focus trials as 'noise'. Responses congruent with subject (agent-like nouns) and object (patient-like nouns) roles were then coded as 'hits' and 'correct rejections', respectively. Incongruent responses were coded as 'misses' (subject focus) and 'false alarms' (object focus).

3. Results

The effect of focus was significant ($\beta = .755$, $SE = .328$, $z = 2.300$, $p = .021$), indicating that participants based their preferred sentence ending on the pitch accent establishing subject or object focus (descriptive statistics in Table 1). The random effects are shown in Table 2.

Table 1: Proportion of responses (in %) congruent with focus condition. SEM = standard error of the mean.

Focus condition	Mean	SEM
Subject	83.1	2.2
Object	70.7	4.6

Table 2: GLMM results (random effects).

Group	Variable	Variance	SD	Corr.
item	intercept	0.374	0.612	
	focus:subj	1.767	1.329	-0.84
subject	intercept	2.741	1.655	
	focus:subj	2.031	1.425	-0.79

The signal detection theory analysis (Figure 3) revealed that most listeners completed the sentences with good sensitivity to the pitch accents in the main clause (high hit rate, low false alarm rate). However, some listeners seemed not to respond according to pitch accents, with a subgroup responding with a bias toward subject-like nouns instead (high hit rate, high false alarm rate).

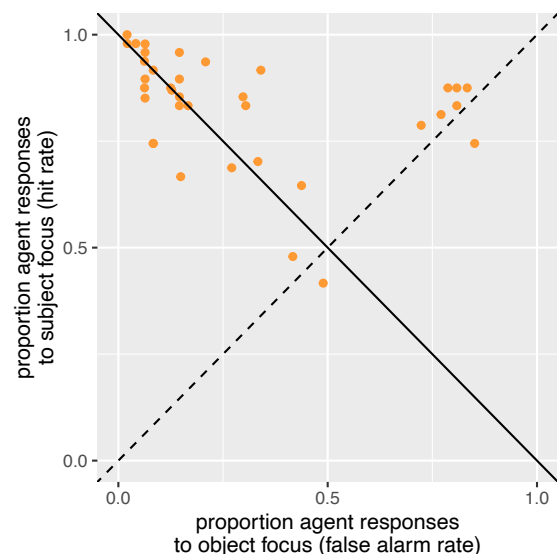


Figure 3: Visualisation of the SDT analysis. Each dot represents an aggregate of individual participant data, plotted as a function of the hit rate and false alarm rate. Data points in the top left corner represent high sensitivity to the prosodic cue (high proportion of agent responses after subject focus, low proportion of agent response after object focus), whereas the top right corner represents a semantic bias towards agent responses (high proportion of agent responses regardless of pitch accent).

4. Discussion

In this study, listeners heard sentences in which either the subject or object was marked by a contrastive pitch accent. The sentences were truncated, omitting the noun phrase of the

sentence-final ellipsis clause. Listeners were asked to complete the sentence with the noun phrase they thought was most suitable. They could complete the sentence with an agent-like noun, congruent with subject focus in the preceding main clause, or a patient-like noun, congruent with object focus. The study had two aims: first, to confirm that pitch accents can establish expectations concerning the subject/object role of an upcoming ellipsis phrase and, second, to explore whether interindividual variability exists in whether listeners base their responses on the available prosodic cues or rather a structural bias.

As to the first aim, group-level results indicated that listeners indeed completed the sentences according to the pitch accents available in the main clause. After a contrastive pitch accent on the subject, listeners preferred an agent-like noun in the ellipsis clause. Conversely, after object focus, they preferred patient-like nouns as sentence continuation. From this main effect of focus a number of conclusions can be drawn. First, listeners were able to perceive the difference in pitch accent patterns. Second, they interpreted these pitch accents to mean that either subject or object was marked with contrastive focus. Third, they derived from this information that the noun phrase in the ellipsis occupied a role that was parallel to the focused constituent. Finally, they were able to indicate this role by choosing the response option that was semantically appropriate for this role (agent or patient). In other words, when implicitly probing the interpretation of focus structure, listeners can give an explicit judgement concerning the semantic features of an upcoming, contrasting constituent. This is in agreement with earlier work showing that listeners can use pitch accents to resolve structural ambiguities [8] and for reference resolution [6]. It extends research on ellipsis resolution, which has pointed toward a role for parallels between noun phrases as well as a role for prosodic cues: while listeners often rely on a semantic or grammatical bias for their interpretation, prosodic cues can partly override these biases and steer the interpretation of focus structure. This is in line with previous work on the effects of prosody using comprehension questions or questionnaires [10], [13]. The current study showed these effects without explicitly probing the interpretation of contrastive focus (albeit by explicitly probing the ellipsis).

The second aim of the study was to explore interindividual differences in the type of cue that listeners base their responses on. In line with the original study, we found considerable interindividual differences in the response patterns. Most listeners demonstrated high sensitivity to the prosodic cues, responding with noun phrases that were semantically congruent with the contrastively accented noun. Yet, a minority of eight participants showed poor sensitivity to the pitch accent pattern, and responded according to a structural bias instead. Specifically, they showed a response preference for agent-like nouns, regardless of the prosody of the antecedent clause. Concerning the apparent lack of sensitivity to prosodic information, a number of explanations can be formulated: listeners may have struggled to detect the prosodic cues, may have been less capable of keeping them in working memory, or perhaps did not manage to derive the appropriate focus structure from the pitch accent pattern. The current design does not allow for a distinction between these possible explanations and further research is needed. While these listeners showed reduced sensitivity to prosody, it remains to be explained why they responded with a bias for agent-like (rather than patient-like) nouns. Since focused material predominantly occurs late in the sentence [26], [27], disregarding prosody would likely

lead to an object-focus interpretation and therefore a patient-noun response. In the original publication, however, an acceptability judgement task on written versions of the materials did not point towards a bias towards object nor subject focus variants of the sentence [16]. Yet, both in the original study (in German) and the current study (in Dutch), participants with reduced sensitivity to prosody favoured agent-like nouns in the ellipsis. A plausible explanation would be that a bare noun in an ellipsis clause, without structure inferred from the antecedent clause, is more likely to attract a subject interpretation and therefore agent-like nouns.

Clearly, a larger sample is required for a thorough individual differences approach, to better understand the nature of this variability. Individual differences can concern domain-general processes and domain-specific ones, pertaining linguistic representations themselves [28]. Considering the current task, a domain-general explanation for the variability in responses could be variability in the sensory response, with some listeners being more perceptive to the acoustic cues than others. Alternatively, variability in working memory capacity could play a role, as participants had to hold the pitch accent pattern in working memory in order to select the appropriate response: suboptimal encoding or retrieval of this memory representation could plausibly lead listeners to rely on a structural bias instead. Domain-specific factors, too, could be at play. These could be individual differences in how acoustic features are mapped onto phonological categories. Furthermore, differences could arise at the level of structure building: some listeners might be more inclined to integrate prosodic information into syntactic structure, while others may rely more on a default interpretation of the antecedent and/or ellipsis clauses. These possible sources of variability could be studied in future research.

5. Acknowledgements

The authors would like to thank Peter Hofmann for help with preparing the stimulus materials and running the online norming experiment. The study was funded by the Max Planck Society.

6. References

- [1] D. R. Ladd and A. Arvaniti, "Prosodic Prominence Across Languages," *Annu Rev Linguist*, vol. 9, no. 1, 2022, doi: 10.1146/annurev-linguistics-031120-101954.
- [2] M. Rooth, "A theory of focus interpretation," *Nat Lang Seman*, vol. 1, no. 1, pp. 75–116, 1992, doi: 10.1007/bf02342617.
- [3] D. Dahan, M. K. Tanenhaus, and C. G. Chambers, "Accent and reference resolution in spoken-language comprehension," *Journal of Memory and Language*, vol. 47, no. 2, pp. 292–314, Aug. 2002, doi: 10.1016/s0749-596x(02)00001-3.
- [4] A. Weber, B. Braun, and M. W. Crocker, "Finding referents in time: eye-tracking evidence for the role of contrastive accents.," *Language and Speech*, vol. 49, no. Pt 3, pp. 367–392, 2006, doi: 10.1177/00238309060490030301.

- [5] D. G. Watson, M. K. Tanenhaus, and C. A. Gunlogson, "Interpreting Pitch Accents in Online Comprehension: H* vs. L+H*," *Cogn Sci*, vol. 32, no. 7, pp. 1232–1244, 2008, doi: 10.1080/03640210802138755.
- [6] M. K. Tanenhaus, M. J. Spivey-Knowlton, K. M. Eberhard, and J. C. Sedivy, "Integration of visual and linguistic information in spoken language comprehension.," *Science*, vol. 268, no. 5217, pp. 1632–1634, Jun. 1995, [Online]. Available: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=7777863&retmode=ref&cmd=prlinks>
- [7] A. Weber, M. Grice, and M. W. Crocker, "The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements.," *Cognition*, vol. 99, no. 2, pp. B63–72, 2006, doi: 10.1016/j.cognition.2005.07.001.
- [8] C. Read, A. Kraak, and L. Boves, "The interpretation of ambiguous who-questions in Dutch: The effect of intonation," in *Linguistics in the Netherlands 1977–1979*, W. Zonneveld and F. Weerman, Eds., 1980. doi: 10.1515/9783110860023-032.
- [9] A. J. Schafer, K. Carlson, C. Clifton, and L. Frazier, "Focus and the Interpretation of Pitch Accent: Disambiguating Embedded Questions.," *Language and Speech*, vol. 43, no. 1, pp. 75–105, 2000, doi: 10.1177/00238309000430010301.
- [10] K. Carlson, *Parallelism and prosody in the processing of ellipsis sentences*. in Routledge. Routledge, 2002.
- [11] S. Winkler, "Ellipsis and Prosody," in *Oxford Handbook of Ellipsis*, J. van Craenenbroeck and T. Temmerman, Eds., in *Oxford Handbook of Ellipsis*, Oxford Handbook of Ellipsis, 2019, pp. 356–386. doi: 10.1093/oxfordhb/9780198712398.013.15.
- [12] B. Stolterfoht, A. D. Friederici, K. Alter, and A. Steube, "Processing focus structure and implicit prosody during reading: differential ERP effects.," *Cognition*, vol. 104, no. 3, pp. 565–590, 2007, doi: 10.1016/j.cognition.2006.08.001.
- [13] L. Frazier, C. Clifton, and K. Carlson, "Focus and VP Ellipsis," *Lang. Speech*, vol. 50, no. 1, pp. 1–21, 2007, doi: 10.1177/00238309070500010101.
- [14] K. Carlson, "Clefting, Parallelism, and Focus in Ellipsis Sentences," in *Grammatical Approaches to Language Processing*, 2nd ed., vol. 46, in *Grammatical Approaches to Language Processing*, vol. 46., Explicit and Implicit Prosody in Sentence Processing, 2015, pp. 63–83. doi: 10.1007/978-3-319-12961-7_4.
- [15] K. Carlson, "How Prosody Influences Sentence Comprehension," *Language and Linguistics Compass*, vol. 3, no. 5, pp. 1188–1200, 2009, doi: 10.1111/j.1749-818x.2009.00150.x.
- [16] C. L. van der Burght, A. D. Friederici, T. Goucha, and G. Hartwigsen, "Pitch accents create dissociable syntactic and semantic expectations during sentence processing.," *Cognition*, vol. 212, p. 104702, 2021, doi: 10.1016/j.cognition.2021.104702.
- [17] X. Xie, A. Buxó-Lugo, and C. Kurumada, "Encoding and decoding of meaning through structured variability in intonational speech prosody.," *Cognition*, vol. 211, p. 104619, Feb. 2021, doi: 10.1016/j.cognition.2021.104619.
- [18] J. Kim, "Individual differences in the production of prosodic boundaries in American English," *J. Acoust. Soc. Am.*, vol. 145, no. 3, pp. 1933–1933, 2019, doi: 10.1121/1.5102039.
- [19] M. Hansen, C. Huttenlauch, C. de Beer, I. Wartenburger, and S. Hanne, "Individual Differences in Early Disambiguation of Prosodic Grouping," *Lang Speech*, p. 002383092211273, 2022, doi: 10.1177/00238309221127374.
- [20] S. Baumann and B. Winter, "What makes a word prominent? Predicting untrained German listeners' perceptual judgments," *Journal of Phonetics*, vol. 70, pp. 20–38, Sep. 2018, doi: 10.1016/j.wocn.2018.05.004.
- [21] K. Kobrock and T. B. Roettger, "Assessing the replication landscape in experimental linguistics," *Glossa Psycholinguist.*, vol. 2, no. 1, 2023, doi: 10.5070/g6011135.
- [22] R. S. Jackendoff, *Semantic Interpretation in Generative Grammar*. 1972. [Online]. Available: <https://eric.ed.gov/?id=ED082548>
- [23] P. Boersma and D. Weenink, *Praat*. 2022. [Online]. Available: <http://www.praat.org/>
- [24] F. Ferreira, "The misinterpretation of noncanonical sentences," *Cogn Psychol*, vol. 47, no. 2, pp. 164–203, Sep. 2003, doi: 10.1016/s0010-0285(03)00005-7.
- [25] N. A. Macmillan and C. D. Creelman, *Detection Theory*. in Cambridge University Press. Cambridge University Press, 1991.
- [26] K. D. Kuthy and B. Stolterfoht, "Focus projection revisited: Pitch accent perception in German," in *Experiments in Focus*, S. Featherston, R. Hörnig, S. von Wietersheim, and S. Winkler, Eds., in *Experiments in Focus*, Experiments in Focus, 2019, pp. 57–70. doi: 10.1515/9783110623093-003.
- [27] K. Carlson, M. W. Dickey, L. Frazier, and C. C. Jr, "Information structure expectations in sentence comprehension," *The Quarterly Journal of Experimental Psychology*, vol. 62, no. 1, pp. 114–139, Jan. 2009, doi: 10.1080/17470210701880171.
- [28] E. Kidd, S. Donnelly, and M. H. Christiansen, "Individual Differences in Language Acquisition and Processing," *Trends in Cognitive Sciences*, vol. 0, no. 0, 2017, doi: 10.1016/j.tics.2017.11.006.