



Understandable Misstatements Lead to Gentle Corrections: Prosodic Realization of Epistemic Gaps

Iris Chuoying Ouyang¹, Elsi Kaiser²

¹ National University of Singapore

² University of Southern California

iris.oy@nus.edu.sg, emkaiser@usc.edu

Abstract

Research has shown that corrective information is produced with higher prosodic prominence than non-corrective information. However, it remains unclear how corrective prosody is realized in different communicative settings. We conducted two production experiments to investigate whether interlocutors' prosodic realization of corrective focus depends on each other's knowledge state. Participants carried out a statement-response task in pairs (e.g. Speaker A: "Tina had shrimp at a restaurant." Speaker B: "No, she had beef at a restaurant."). We manipulated whether Speaker A's statement was implausible in the context (e.g. Tina in fact hates seafood). Furthermore, the two experiments differed in whether Speaker A knew the probability of his or her statement. In Exp.1, both speakers had access to the crucial context concerning the probability of Speaker A's statement (e.g. Tina's preferences about food). In Exp.2, only Speaker B had access to this background information. Mixed-effects models were fit on the f0 ranges of target words in Speaker B's responses (e.g. "beef"). We found that Speaker B's prosody was influenced by both (i) the contextual probability of Speaker A's statements and (ii) Speaker A's knowledge (or lack thereof) about the contextual probability. We present an analysis where the prosodic prominence associated with corrective information reflects the gap between expectation and reality – in this case, what Speaker B had expected Speaker A to know and what Speaker A appeared to know.

Index Terms: corrective focus, addressee's knowledge state, contextual probability, epistemic surprisal, f0 ranges

1. Introduction

It is widely accepted that prosodic prominence can signal the extent to which a linguistic element is 'informative'. The acoustic properties of an utterance such as duration, f0, intensity, and spectral characteristics provide cues for the relative informativity of its components (see [1] for a review). Prior work has approached the relationship between prosody and informativity from various angles. Two of the popular ones are **information structure** and **perspective-taking**:

In the information-structural tradition of research, acoustic prominence is associated with linguistic material in the foreground – broadly speaking, material that adds new information to the conversation. Depending on the preceding discourse, speakers may prosodically emphasize particular words in an utterance to direct their addressee's attention to the important message they are trying to convey. One type of information structure that has been extensively studied is

contrastive focus, of which various subtypes have been identified (e.g. [2]). In this paper, we concentrate on a subtype of contrastive focus that has been referred to as *corrective focus* (e.g. [3]), because its information-structural properties are well-understood and it is prevalent in communication. Corrective focus conveys messages contradicting to information that is already present in the discourse (see ex. 1). Contrastive/corrective elements have been shown to receive greater acoustic prominence than non-contrastive/non-corrective elements (e.g. [4-8]).

Furthermore, as language communication involves not only the talker/speaker but also the listener/addressee, the informativity of a word could depend on the addressee's knowledge about the word. Because people often have different knowledge backgrounds, this brings up the question of whether speakers' encoding of information is addressee-oriented, taking into account the audience's knowledge state, or whether it is egocentric, driven by speaker-internal considerations. The general question of perspective-taking has received considerable attention in the literature (e.g. [9-18]). Research shows that speakers may vary their prosody based on their addressee's knowledge state, although a consensus has not been reached on the cognitive processes and mechanisms behind this phenomenon.

Evidence suggests that perspective-taking involves at least two dimensions: (i) what speakers have assumed about their interlocutors at the start of the conversation (e.g. [19-23]), and (ii) what speakers learn about their interlocutors during the conversation, based on their interlocutors' utterances or behavior (e.g. [24-25]). However, little attention has been paid to the interplay between these two factors (but see [26-28]). Moreover, it remains unclear how perspective-taking factors interact with information structure in shaping the prosody of utterances, since few studies have investigated this question in general (but see [11] and [14]).

1.1. Aims and Predictions

As discussed above, research has shown that the prosodic representation of an utterance depends on how informative its constituents are relative to each other. Information structure, such as corrective focus, and perspective-taking factors, such as the speaker's prior expectations and the addressee's actual behavior, all play a role in the prosodic encoding of informativity. However, little work has been done on the potential interaction between these factors. To shed light on this issue, we conducted a psycholinguistic production study that investigated whether perspective-taking factors could affect the prosodic prominence associated with information-structural focus. Specifically, the structure of corrective focus

essentially comprises an incorrect message and a corrective response. In daily conversation, these two components of corrective focus structure are often spoken by different conversational participants – would the prosody of the corrective response be affected by the speaker’s perception of the other person? For example, if the other person’s mistake is unexpected, would the speaker increase or attenuate the prosodic prominence s/he produces in his/her corrective response? What if the other person’s mistake is expected?

In terms of the acoustic correlate of prosodic prominence, we concentrate on the size of excursions in an f0 contour (which will be called ‘f0 range’ henceforth) in this paper. We chose f0 because it is an acoustic dimension that has been extensively studied in the information-structural tradition yet not much so in the perspective-taking tradition. In other words, by conducting this study, we also hoped to provide further evidence for the effects of perspective-taking factors on f0. It has been found that information-structural focus results in greater f0 protrusion, increased f0, or decreased f0 on the focused element (e.g. [4] [6], [7] and [28]). Therefore, we quantitatively measured f0 ranges, which presumably would capture the level of prominence in the f0 dimension. Based on these existing studies, we predicted that corrective responses would have larger f0 ranges than non-corrective responses.

If our prediction for corrective focus is borne out, we can then look into how speakers’ prior expectations and their interlocutors’ actual behavior interact in shaping corrective prosody. We expected to see effects of both factors, based on the perspective-taking research discussed in the preceding section.

2. Methods

We conducted two production experiments (Exp 1 and 2) with interactive setups. Each trial consisted of a production task and a subsequent comprehension task. Naïve participants worked in pairs on the production task and independently on the comprehension task. The production task provided the critical recording, namely the target sentence in the dialogue. The comprehension task was included to engage participants in the production task, as paying attention to the dialogue was necessary to successfully perform the comprehension task. (We do not discuss the comprehension task in detail here because it is not relevant for the results, but participants essentially had to answer a *wh*-question about the dialogue.

2.1. Design and Procedures

Two naïve participants (Speakers A and B) worked with each other in reading aloud dialogues. In both experiments, the primary speaker of interest is Speaker A. We will first present the setup of Exp 1, and then turn to Exp 2.

In Exp 1, each dialogue consisted of five sentences (S1-5), as shown in examples (1-2). Participants saw the text of the sentences on paper. S1 and S2 were spoken by Speaker A, introducing a character (e.g. Jacky) and his or her preference or need (e.g. Jacky likes fruit in her salads but not vegetables, or Zac likes vegetables in his salads but not fruit). S3 and S4 were spoken by Speaker B. S3 provided information about a recent event (e.g. Jacky/Zac going grocery shopping). S4 commented on this event, starting with “I heard that...” and describing something this person did. Crucially, this described event was either plausible or implausible in the context of S1-2 (e.g. plausible: Jacky got apples; implausible: Zac got

apples). Once S4 was produced, the other speaker, Speaker A, spoke S5, starting with either “Yes” or “No” to confirm or correct the previous sentence. Participants thus interacted with each other to produce the dialogues, and each participant only had access to the text of sentences that he or she was responsible for.

CORRECTIVE FOCUS (1)

A: Jacky prefers her salad a certain way. [S1]
 A: She loves fruit but hates vegetables. [S2]
 B: She went grocery shopping yesterday evening. [S3]
 B: I heard that she got some **apples** PROBABLE at the farmer’s market. [S4]
 A: No, she got some **[lettuce]** CORRECTIVE at the farmer’s market. [S5]

NON-CORRECTIVE INFORMATION (2)

A: Zac tends to put certain things in his salad. [S1]
 A: He loves vegetables but hates fruit. [S2]
 B: He went grocery shopping this morning. [S3]
 B: I heard that he got some **apples** IMPROBABLE at the supermarket. [S4]
 A: Yes, he got some **[apples]** NON-CORRECTIVE at the supermarket. [S5]

The experiment had 192 target items and 96 fillers. Each pair of participants encountered 48-96 items and did not see any item more than once. Each participant served as Speaker A in half of the dialogues and Speaker B in the other half, i.e., the roles of speaker A and B were intermixed throughout the experiment. These two halves had different sets of characters and scenarios. For example, one participant began and finished all the dialogues that involved Gary, Lauren, and the part of their house that they need to buy new things for (e.g. bathroom vs. patio), whereas the other participant began and finished all the dialogues about Jacky, Zac, and the kind of salad they like (e.g. fruit vs. vegetable).

In all dialogues, S4 (Speaker B’s statement) and S5 (Speaker A’s response) were the critical sentences, which contained transitive clauses with the following structure: a third-person singular pronoun subject, a simple past tense verb, an object noun phrase, and a prepositional phrase indicating a location or beneficiary. The critical word of interest is the head noun of the object noun phrase (e.g. apples or lettuce). To investigate how the addressee’s knowledge state affects the speaker’s corrective prosody, we manipulated the information-structural status of Speaker A’s response with respect to Speaker B’s statement, as exemplified in (1-2), and the contextual probability of Speaker B’s statement, relative to the character’s preference and need. The contextual probability of Speaker A’s responses was counterbalanced (except on non-corrective trials, where the object noun was the same between B’s statement and A’s response). These contextual probabilities were estimated through a web-based norming study (conducted on Amazon Mechanical Turk, <https://www.mturk.com/>).

Furthermore, in order to distinguish what people initially expect about their interlocutor from what they learn about their interlocutor during the conversation, we conducted a parallel experiment, Exp 2, and manipulated Speaker B’s knowledge state between the two experiments. Specifically, in Exp 2, we created a knowledge gap between Speakers A and B by removing Speaker B’s access to S2, the sentence introducing

the character's preference/need and thereby establishing a contextual bias. Participants were asked not to read S2 aloud, which essentially made it Speaker A's private knowledge. In other words, the contextual probability of the critical words was privileged information kept by Speaker A and unavailable to Speaker B. Everything else was held the same between the experiments.

Thus, three independent variables were implemented:

(i) **Correctiveness of the response (Corrective vs. Non-Corrective)**, whether the object head noun in Speaker A's response (S5) was in corrective focus. For instance, in ex. 1, the object 'lettuce' in Speaker A's response (*No, she got some lettuce at the farmer's market*) is corrective information because Speaker A was correcting Speaker B's statement (S4, *I heard that she got some apples at the farmer's market*). In contrast, in ex. 2, the object 'apples' in Speaker A's response (*Yes, he got some apples at the farmer's market*) is non-corrective information because Speaker A was confirming Speaker B's statement (*I heard that he got some apples at the supermarket*).

(ii) **Contextual probability of the statement; Statement Type (Probable vs. Improbable)**: whether the object head noun in Speaker B's statement (e.g. *apples* in ex. 1) matched or conflicted with the character's preference/need. For instance, in ex. 1, Speaker B says that s/he heard Jacky bought apples, and we know Jacky loves fruit, so ex. 1 shows a contextually probable statement. However, consider an alternative where Speaker B says that s/he heard Jacky bought lettuce. Because we know from the preceding context that Jacky hates vegetables, 'she bought lettuce' is a contextually improbable statement here.

(iii) **Knowledge Type (Shared vs. Privileged)**, whether the contextual probability of the object head nouns was shared knowledge between the interlocutors (Exp 1) or Speaker A's privileged knowledge (Exp 2).

The dependent variable we measured was the f0 range (calculated by subtracting the f0 minimum from the f0 maximum) in the object head noun of Speaker A's response (e.g. *lettuce* in ex. 1).

2.2. Participants

44 self-claimed native speakers of American English participated, 11 pairs in each experiment. One participant was excluded from the data analysis because of a non-native accent. All participants were students or staff at the University of Southern California.

2.3. Data Analysis

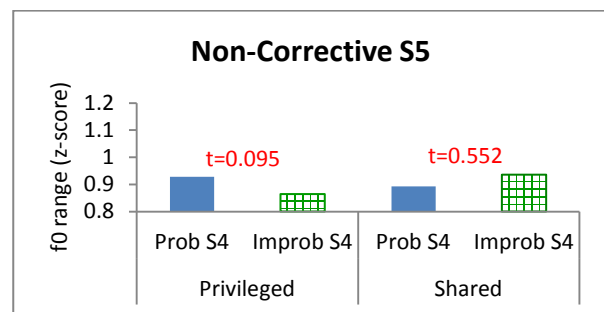
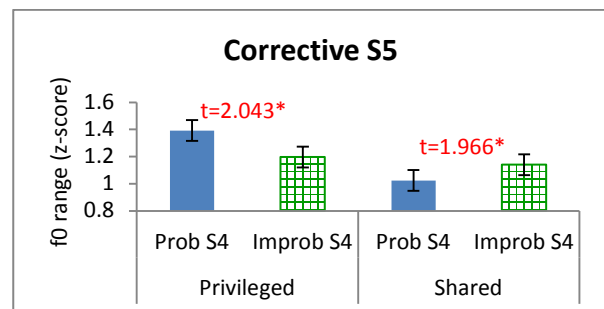
We analyzed Speaker A's responses (S5), as mentioned in Section 2.1. 1632 utterances were collected from the 43 participants, each producing 24-48 target responses. Out of the full set of data, 171 utterances (10.5%) were not included in the data analysis due to speech errors, disfluencies, or technical issues with the audio recording.

F0 measurements were obtained using the YAAPT (Yet Another Algorithm for Pitch Tracking) algorithm [35]. The raw f0 values were then smoothed (*smoothn* in MATLAB: [36]) to remove f0 tracking errors and segmental effects. The smoothed values were then converted into a semitone scale, as semitones reflect pitch perception better than the Hertz scale (e.g. [37]). Finally, the data were normalized by subject using

z-scores, to factor out individual differences in f0 registers (e.g. women usually have wider and higher registers of f0 than men). The z-scores represented each data point in terms of its number of standard deviations above or below the mean across all utterances produced by a given speaker.

Mixed-effects models were conducted on f0 ranges (*anova* in R: [38]; *lme4* in R: [39]; *lmerTest* in R: [40]). Correctiveness, Statement Type, and Knowledge Type were included as fixed effects; Subject and Scenario were included as random effects. When specifying the structure of random effects, we started with a full model (i.e. including intercepts and slopes for Subject and Scenario) and excluded a random slope when it did not significantly contribute to the model. The final analysis did not have random slopes but included random intercepts for Subject and Scenario.

3. Results



Figures 1-2: Mean f0 ranges of Speaker A's responses (S5) in each condition.

Overall, the predictions outlined in Section 1.1 were borne out. Correctiveness had a main effect ($t = 2.919$, $p < 0.01$) and no interaction with either Statement Type ($t = 1.082$, $p = 0.280$) or Knowledge Type ($t = 0.744$, $p = 0.457$). Speaker A's corrective responses had significantly larger f0 ranges than their non-corrective responses across the board (corrective mean = 1.168; non-corrective mean = 0.908). Furthermore, Statement Type and Knowledge Type interacted with each other ($t = 2.827$, $p < 0.01$). Although the three-way interaction between Correctiveness, Statement Type and Knowledge Type was not found significant ($t = 1.170$, $p = 0.242$), planned comparison revealed that the interaction between Statement Type and Knowledge Type affected f0 ranges in the corrective responses, but not in the non-corrective responses¹:

¹ Here we primarily focused on the differences between Statement Types (probable vs. improbable) and did not

When the contextual probability of the object nouns was shared knowledge, Speaker B's improbable misstatements elicited significantly larger f0 ranges (in Speaker A's corrective responses) than their probable misstatements ($t = 1.966$, $p < 0.05$; improbable mean = 1.140; probable mean = 1.024). In contrast, when the contextual probability of the object nouns was Speaker A's privileged knowledge, Speaker B's probable misstatements elicited significantly larger f0 ranges (in Speaker A's corrective responses) than their improbable misstatements ($t = 2.043$, $p < 0.05$; probable mean = 1.392; improbable mean = 1.196). Such interaction was not found in non-corrective responses (Statement Type in Exp 1: $t = 0.552$, $p = 0.581$; Statement Type in Exp 2: $t = 0.095$, $p = 0.924$). These patterns can be seen in Figures 1-2.

4. Discussion

This study investigated the modulation of corrective prosody based on the speaker's expectation and realization about the addressee's knowledge state. As discussed in Section 1, previous research has examined prosody from these angles but has not paid close attention to their interplay. A better understanding of these issues is important because they are concerned with fundamental questions regarding the communicative functions of prosody.

Our results show that corrective prosody is influenced both by (i) speakers' prior expectations about their addressees and by (ii) what speakers learn about their addressees through the conversation. Our participants consistently emphasized corrective information over non-corrective information in terms of f0 ranges. However, the degree of emphasis they placed on corrective information varied depending on their conversational partner's knowledge state. In Exp 1, where both participants were expected to be fully aware of the context, they corrected each other with greater prosodic prominence when the misstatements were contextually improbable. In Exp 2, where the speaker of the misstatements (Speaker B) was expected to have no knowledge about the context, the speaker of the corrections (Speaker A) produced greater prosodic prominence when the misstatements were contextually probable. This supports our general prediction that *corrective prosody is modulated by the addressee's knowledge state through the interplay between the speaker's prior expectations and the addressee's actual behavior*.

Why should the speaker's prior expectations and the addressee's actual behavior interact and influence corrective prosody in the way we observed? We suggest that our results can be explained in terms of **the epistemic gap between expectation and reality**. The level of prosodic prominence, such as f0 ranges, might reflect the extent to which speakers are surprised at what they encounter during the conversation. In the case of this study, participants had assumptions about what each other knew, and when the other person's utterances contradicted those prior assumptions, they marked this surprisal prosodically in their responses:

Let us consider the communicative setting in Exp 1, where the characters' preferences and needs were mentioned in the dialogues. Expecting their partner to have this knowledge, participants might not have found it surprising

directly compare between Knowledge Types (shared vs. privileged), because Knowledge Type was a between-subjects variable.

when their partner had misbeliefs consistent with the characters' preferences and needs. In contrast, it might have struck participants as surprising when their supposedly-informed partner had misbeliefs conflicting with the context. This epistemic gap might be what elicited extra prosodic prominence in the corrective responses to contextually improbable misstatements.

Probably for similar reasons, the opposite patterns were yielded by the communicative setting in Exp 2, where the characters' preferences and needs were privy to the responder. Here, participants might have instead been surprised at their partner's probable misstatements, because it might have seemed that their partner was somehow able to form probable beliefs without the critical knowledge. In contrast, improbable misstatements might not have struck participants as surprising, since improbable beliefs might have fit their assumption about their partner's lack of critical knowledge. In other words, the extra prosodic prominence still reflects an epistemic gap, but compared to Exp 1, (im)probability of the misstatements had the opposite effect in Exp 2, because the participants had the opposite expectation about their partner's knowledge state.

Our findings thus add to the small body of literature concerning how speakers update their assumptions about their addressee by taking incoming cues from their addressee's behavior (e.g. [25-27]). Existing studies have shown that speakers' responses are influenced by both what they assume their addressee to know and what their addressee appears to know. This study shows that the prosody of speakers' responses can reveal the gap between speakers' expectations (e.g. what other people are likely to say) and the reality they encounter (e.g. what other people actually say). In the research on prosody and information structure, this study draws attention to the importance of investigating the conversational partner's utterance (e.g. the misstatement in corrective focus structure) that elicits the speaker's response (e.g. the correction). This is consistent with views that highlight the role of interlocutors in language communication (see Sections 1 for more details).

Our idea that speakers' expectations modulate the encoding of corrective information is in line with work in other areas of experimental linguistics. Indeed, a substantial amount of research has investigated the effects of expectations in language comprehension [35-42] and production [43-47]. Predictive processing has been shown to occur at various linguistic levels, and our work contributes to this literature by exploring the prosodic domain.

5. Acknowledgements

We gratefully acknowledge the support of a Doctoral Dissertation Research Improvement grant from the National Science Foundation.

6. References

- [1] Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and cognitive processes*, 25(7-9), 905-945.
- [2] Vallduví, E., & Vilkuna, M. (1998). On rheme and kontrast. *Syntax and semantics*, 79-108.
- [3] Dik, S. C., & Hengeveld, K. (1997). *The theory of functional grammar: the structure of the clause*. Walter de Gruyter.
- [4] Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and cognitive processes*, 25(7-9), 1044-1098.

- [5] Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *The Journal of the acoustical society of America*, 77(6), 2142-2156.
- [6] Couper-Kuhlen, E. (1984). A new look at contrastive intonation. In *Modes of interpretation: Essays presented to Ernst Leisi* (pp. 137-158). Gunter Narr Verlag.
- [7] Katz, J., & Selkirk, E. (2011). Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language*, 87(4), 771-816.
- [8] Krahmer, E., & Swerts, M. (2001). On the alleged existence of contrastive accents. *Speech communication*, 34(4), 391-405.
- [9] Arnold, J. E., Kahn, J. M., & Pancani, G. C. (2012). Audience design affects acoustic reduction via production facilitation. *Psychonomic bulletin & review*, 19(3), 505-512.
- [10] Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of memory and language*, 42(1), 1-22.
- [11] Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee?. *Journal of memory and language*, 62(1), 35-51.
- [12] Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground?. *Cognition*, 59(1), 91-117.
- [13] Kahn, J. M., & Arnold, J. E. (2012). A processing-centered look at the contribution of givenness to durational reduction. *Journal of memory and language*, 67(3), 311-325.
- [14] Kaland, C., Swerts, M., & Krahmer, E. (2013). Accounting for the listener: Comparing the production of contrastive intonation in typically-developing speakers and speakers with autism. *The Journal of the acoustical society of America*, 134(3), 2182-2196.
- [15] Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic bulletin & review*, 9(3), 550-557.
- [16] Roßnagel, C. (2000). Cognitive load and perspective-taking: applying the automatic-controlled distinction to verbal communication. *European journal of social psychology*, 30(3), 429-445.
- [17] Roßnagel, C. S. (2004). Lost in thought: Cognitive load and the processing of addressees' feedback in verbal communication. *Experimental psychology*, 51(3), 191-200.
- [18] Rosa, E. C., & Arnold, J. E. (2011). The role of attention in choice of referring expressions. *Proceedings of PRE-Cogsci: Bridging the gap between computational, empirical and theoretical approaches to reference*.
- [19] Baker, R. E., & Bradlow, A. R. (2009). Variability in word duration as a function of probability, speech style, and prosody. *Language and speech*, 52(4), 391-413.
- [20] Ferguson, S. H. (2004). Talker differences in clear and conversational speech: Vowel intelligibility for normal hearing listeners. *Journal of the acoustical society of America* 116:2365-2373.
- [21] Ferguson, S. H., & Kewley-Port, D. (2007). Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of speech, language, and hearing research* 50:1241-1255.
- [22] Fussell, S. R., & Krauss, R. M. (1989a). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of experimental social psychology*, 25(3), 203-219.
- [23] Hummert, M. L., & Shaner, J. L. (1994). Patronizing speech to the elderly as a function of stereotyping. *Communication studies*, 45(2), 145-158.
- [24] Isaacs, E. A., & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of experimental psychology: general*, 116(1), 26.
- [25] Brennan, S. E. (1991). Conversation with and through computers. *User modeling and user-adapted interaction*, 1(1), 67-86.
- [26] Kuhlen, A. K., & Brennan, S. E. (2010). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes*, 47(7), 567-587.
- [27] Russell, A. W., & Schober, M. F. (1999). How beliefs about a partner's goals affect referring in goal-discrepant conversations. *Discourse processes*, 27(1), 1-33.
- [28] Brown, G. (1983). Prosodic structure and the given/new distinction. In *Prosody: Models and measurements* (pp. 67-77). Springer Berlin Heidelberg.
- [29] Zahorian, S. A., & Hu, H. (2008). A spectral/temporal method for robust fundamental frequency tracking. *The Journal of the Acoustical Society of America*, 123(6), 4559-4571.
- [30] Garcia, D. (2010). Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational statistics & data analysis*, 54(4), 1167-1178.
- [31] Nolan, F. (2003). Intonational equivalence: an experimental evaluation of pitch scales. In *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona* (Vol. 39).
- [32] Chambers, J. M. and Hastie, T. J. (1992) *Statistical Models in S*, Wadsworth & Brooks/Cole.
- [33] Bates, Douglas, Martin Maechler, Ben Bolker, and Steven Walker. 2014. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7.
- [34] Kuznetsova, Alexandra, Per Bruun Brockhoff, and Rune Haubo Bojesen Christensen. 2015. lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-25.
- [35] Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247-264.
- [36] Boland, J. E. (2005). Visual arguments. *Cognition*, 95(3), 237-274.
- [37] Brown, M., Salverda, A. P., Dilley, L. C., & Tanenhaus, M. K. (2011). Expectations from preceding prosody influence segmentation in online sentence processing. *Psychonomic bulletin & review*, 18(6), 1189-1196.
- [38] Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of memory and language*, 58(2), 541-573.
- [39] Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive science*, 32(7), 1232-1244.
- [40] Van Berkum, J. J., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of experimental psychology: learning, memory, and cognition*, 31(3), 443.
- [41] DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature neuroscience*, 8(8), 1117-1121.
- [42] Wicha, N. Y., Bates, E. A., Moreno, E. M., & Kutas, M. (2003). Potato not Pope: human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience letters*, 346(3), 165-168.
- [43] Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the acoustical society of America*, 113(2), 1001-1024.
- [44] Clopper, C. G., & Pierrehumbert, J. B. (2008). Effects of semantic predictability and regional dialect on vowel space reduction. *The Journal of the acoustical society of America*, 124(3), 1682-1688.
- [45] Pan, S., & Hirschberg, J. (2000). Modeling local context for pitch accent prediction. In *Proceedings of the 38th annual meeting on association for computational linguistics* (pp. 233-240). Association for Computational Linguistics.
- [46] Scarborough, R. (2010). Lexical and contextual predictability: Confluent effects on the production of vowels. *Laboratory phonology*, 10, 557-586.
- [47] Van Son, R. J. J. H., & Pols, L. C. (1999). An acoustic description of consonant reduction. *Speech communication*, 28(2), 125-140.