Speech Prosody 2024
2-5 July 2024, Leiden, The Netherlands

Does communicative skill predict individual variability in the prosodic encoding of lexical and referential givenness?

Janne Lorenzen, Stefan Baumann

I*f*L Phonetik, University of Cologne, Germany
janne.lorenzen@uni-koeln.de, stefan.baumann@uni-koeln.de

Abstract

We investigated individual variability in the prosodic encoding of lexical and referential givenness in German. Additionally, we related this variability to self-assessed communicative skill in our participants. In an interactive reading task, we collected data from 20 speakers producing eight short stories. In each story, the same target word occurred as lexically (*l-*) and referentially (*r-*)*new* or *given*, and in combinations of these levels. We measured several prosodic correlates of prominence in the target words.

Across speakers, *l-new* referents were marked by longer duration and higher periodic energy than *l-given* referents, while *r-new* referents were marked by higher periodic energy and higher intensity than *r-given* ones. While speakers were remarkably similar in their encoding of *l-givenness*, only differing in how strongly they modified duration and periodic energy, there was a more striking contrast in the encoding of *r-givenness*: One group of speakers exclusively relied on periodic energy and intensity, the other group additionally used higher F0 to mark *r-newness*. Differences in the produced givenness contrasts across speakers proved to be related to communicative skill, albeit in opposite ways: *L-givenness* was marked more strongly by speakers with higher communicative skill, *r-givenness* was marked more strongly by speakers with lower communicative skill.

Index Terms: prominence, German, information status, givenness, individual variability, communicative skill

1. Introduction

In this paper, we investigate inter-individual variability in the prosodic encoding of lexical and referential givenness and its relation to communicative skill. Following [1], givenness (or information status) can be understood as the degree of cognitive activation of a referent determined by the discourse context. *New* referents, which have not been established in the discourse, are regarded as cognitively inactive, *given* referents, which are already well-established, are cognitively active.

In many West Germanic languages including German, givenness is encoded via prosodic prominence in that *new* referents are produced with longer duration, higher intensity and more extensive F0 movement than *given* referents [2-5]. Two dimensions of givenness have been distinguished that have an influence on prosodic prominence: lexical and referential givenness [4]. Lexical givenness relates to the number of and distance between mentions of a lexical item. When a lexical item is mentioned for the first time, it can be regarded a lexically *new (l-new)*, in any repeated mention (within the next five intonation phrases) it is lexically *given (l-given)*, no matter if it is coreferential with the first mention or not [4, p. 24]. Referential givenness is determined by coreference. If an entity is first introduced to the discourse, it is referentially *new (r-new)*, any following entities that are coreferential with their antecedent are referentially *given (r-given)*. Note that we do not consider focus, another information structural construct that is orthogonal to givenness and may thus override givenness effects, especially in contrastive contexts.

Many previous studies have noted substantial individual variability in the production of prosodic prominence. For example, [6] and [7] observed that German speakers differed in which prosodic parameters they used to mark focus types and how clearly they distinguished between these categories. Similarly, [8] found that speakers of American English differed in how they employed pitch contours to encode informativity.

Although several studies have taken note of inter-individual differences, few production studies investigate reasons behind this variability. Previous perception experiments, however, have found a correlation between the processing of prosodic prominence and communicative skill (e.g., [9, 10]). In [9], American English listeners with higher communicative skill were found to perform better in a prominence rating task than listeners with lower communicative skill. Similarly, in [10], listeners with higher communicative skill more successfully distinguished between H* and L+H* accents. These studies thus assume that listeners with lower communicative skill are less sensitive to the prosody-meaning mapping. We will explore whether such an effect can also be found in the *production* of prosodic prominence.

Attempts to relate communicative skill to prosody production are less frequent and have so far yielded mixed results. [11] found that American English speakers with higher communicative skill produced more acoustically distinct vowel categories than speakers with lower communicative skill. However, this was only true for female speakers. In contrast, [12] observed that American English speakers with *lower* communicative skill more strongly encoded semantic-pragmatic predictability prosodically, while second mention reduction was not influenced by communicative skill in any way. [12] interprets this finding as evidence against lister-oriented theories of speech production.

In the present study, we ask how speakers differ in their prosodic marking of lexical and referential givenness and whether any observed differences relate to communicative skill. To address this issue, we collected production data in a reading task along with information on the communicative skill of our participants using a subset of the Autism-Spectrum-Quotient questionnaire [13], i.e., ten questions asking participants to self-rate their communicative skill. Following findings in previous perception studies, we expect speakers with higher communicative skill to show a clearer mapping of prosody to meaning, i.e., to produce larger prosodic contrasts between both *l-new* and *l-given* and *r-new* and *r-given* target words.

10.21437/SpeechProsody.2024-162

2. Methods

2.1. Reading material

We collected production data via an interactive reading task. Reading material consisted of eight short stories, each containing one target word in five different conditions (see (1) for an example, translated to English). Here, we focus on a subset of this data, i.e., the three conditions where the target word occurs in pre-final position (marked in bold in (1), see i) to iii) for the full target sentences in German). These target words are all grammatical objects with penultimate stress that differ in lexical and referential givenness with the target word in (i) being *l-new* and *r-new*, in (ii) being *l-given* and *r-given* and in (iii) being *l-given* but *r-new*. The target words in (i) and (iii) are marked by indefinite articles to indicate referential *newness*, while the *r-given* target word in (ii) is marked as definite.

(1) *Susanne has been working as an animal keeper at the zoo for many years. She is very happy there and lovingly looks after all her charges. Today Susanne brought **a banana** (i) with her. She had brought a big box with lots of fruit and vegetables, and especially the monkeys were very excited. Then Susanne took out **the banana** (ii). The monkeys immediately gathered around her. The banana looked really tasty and little monkey Tobi desperately wanted to have it. Just a few days ago, a girl and her parents had been standing in front of the monkey enclosure watching Tobi. In doing so, the girl ate **a banana** (iii). Tobi could only look at the beautiful yellow fruit, but now he finally wants to try one! He will definitely enjoy the banana a lot.*

(i) Heute hat Susanne **eine Banane** mitgebracht.

(ii) Dann hat Susanne **die Banane** herausgeholt.

(iii) Dabei hat das Mädchen **eine Banane** gegessen.

2.2. Participants and procedure

We recorded 20 native speakers of German. Participants were between 20 and 31 years old, 13 identified as female, seven as male. They gave informed consent to participate in this experiment and received a monetary compensation. Participants performed an interactive reading task, where a listener was present, who completed three comprehension questions after each story had been read to them. In addition, participants rated their communicative skill on ten scales (i.e., a subset of the questionnaire designed by [13]). Participants could receive up to one point per scale for an overall score of 0 to 10 points, where 0 indicates very low and 10 very high communicative skill. The participants produced 480 utterances. One speaker was excluded from the analysis for failure to complete the questionnaire. We further excluded nine utterances due to hesitations or slips of the tongue, so that 449 utterances entered the analysis.

2.3. Measurements

We collected several continuous prosodic measures, i.e., duration and intensity, periodic energy mass and Delta F0 from the ProPer toolbox [14] and Tonal Center of Gravity (TCoG) scaling [15, 16]. The continuous measures may also reflect categorical contrasts, such as different accentuation patterns or pitch accent types, which we do not explicitly consider here (but see [17] for an analysis using prosodic categories in a subset of the data). More specifically, we measured the duration of the stressed syllable of the target words in milliseconds. Intensity was measured as the mean of the intensity curve in the stressed syllable in dB. Periodic energy mass was measured as the area under the periodic energy curve and represents the integral of duration and intensity in the sonorant portions of the stressed syllable. To capture Delta F0, the Center of Mass (CoM) was determined in the stressed and its preceding syllable. The CoM is a point in time that splits the area under the periodic energy curve into two equally large parts. F0 was then measured at the CoM of these two syllables. The difference in semitones is the Delta F0, which captures a local change in F0. Lastly, we measured the average weighted F0 over the target word and calculated TCoG scaling in semitones relative to a general baseline of 75 Hz for male speakers and 120 Hz for female speakers. TCoG scaling thus constitutes a more global representation of F0 height. All parameters were z-scored before entering the statistical analysis. The communicative skill variable was centered by subtracting the mean from each value.

2.4. Statistical analysis

We ran five Bayesian linear mixed-effects models, one for each prosodic parameter. Each model contained lexical and referential givenness as well as communicative skill as predictors and the interaction of the givenness variables with communicative skill. In addition, we included random intercepts for speaker and item and the by-speaker and by-item random slopes for lexical and referential givenness. We consider the effect of a predictor variable as reliable if the 90% credible interval (CI) does not include 0 and the probability that $\beta>0$ is larger than 0.95. To identify different strategies employed by groups of speakers, we conducted cluster analyses on the random by-speaker slopes extracted from the models following [18]. Models were run in R [19] using the *brms* package [20] as an interface to the *Stan* modeling language [21]. Data and analysis script are available on OSF (https://osf.io/ev6zm/).

3. Results

3.1. Overall results

First, we will consider the results of all speakers taken together. Table 1 summarizes the model parameters describing the effects of lexical givenness on the measured prosodic cues. Due to the centering of the communicative score variable, these results are interpretable with regard to the mean communicative score, which is most representative for our sample. Following the criteria for reliability specified in section 2.4, differences in syllable duration and periodic energy mass are reliably predicted by lexical givenness, in that *l-new* referents are produced with longer duration and higher periodic energy mass than *l-given* referents. *L-new* referents are also generally produced with higher intensity, Delta F0 and TCoG scaling values compared to *l-given* referents, although these differences are not reliable.

Table 1: *Model parameters describing the effects of lexical givenness.*

Variable	β	90% CI	Pr(β>0)
Syllable duration	0.44	[0.21; 0.67]	0.99
Intensity	0.13	[0.00; 0.26]	0.94
Periodic energy mass	0.37	[0.20; 0.54]	1.00
Delta F0	0.09	[-0.16; 0.34]	0.75
TCoG scaling	0.09	[-0.05; 0.23]	0.87

Table 2 summarizes the model parameters describing the effects of referential givenness. *R-new* referents are reliably marked by higher intensity and periodic energy mass compared to *r-given* referents. There are also tendencies for *r-new* referents to be produced with higher Delta F0 and TCoG scaling than *r-given* referents. Syllable duration, however, does not seem to encode referential givenness.

Table 2: *Model parameters describing the effects of referential givenness.*

Variable	β	90% CI	Pr(β>0)
Syllable duration	-0.01	[-0.23; 0.21]	0.47
Intensity	0.25	[0.11; 0.39]	1.00
Periodic energy mass	0.42	[0.24; 0.60]	1.00
Delta F0	0.20	[-0.09; 0.49]	0.87
TCoG scaling	0.14	[-0.02; 0.30]	0.93

3.2. Individual strategies

Next, we focus on the by-speaker random slopes to identify potential different strategies for the encoding of lexical and referential givenness.

For lexical givenness, two clusters of speakers emerge from the analysis: Cluster 1 contains 12 speakers and Cluster 2 contains 7 speakers. Figure 1 shows the average random slopes values with standard errors for the effect of lexical givenness on the prosodic parameters per speaker group. Here, a positive value indicates that *new* referents have a higher value than *given* referents in the parameter in question. Both clusters are remarkably similar, especially in their (moderate) use of higher Delta F0, TCoG scaling and intensity. The main difference between these speakers is the strength of encoding via periodic energy mass and syllable duration, in that Cluster 2 makes a stronger difference between *l-new* and *l-given* referents in terms of these two parameters.

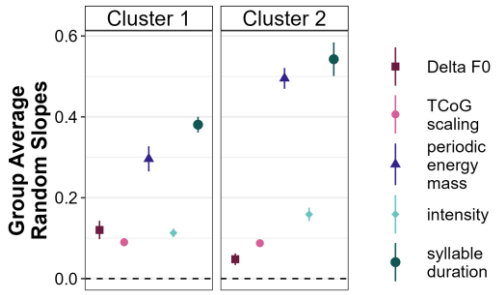

Figure 1: *Results of cluster analysis for lexical givenness.*

Considering referential givenness, again two clusters of speakers emerge. Cluster 1 contains 14 speakers while Cluster 2 only contains 5 speakers. Figure 2 illustrates the characteristics of the two groups. Here, Cluster 1 marks *r-new* referents with higher periodic energy mass and intensity. Delta F0, TCoG scaling and syllable duration are not used to differentiate *r-new* and *r-given* referents by these speakers. While Cluster 2 is similar to Cluster 1 in their use of periodic energy mass, intensity and syllable duration, this cluster additionally produces higher Delta F0 and higher TCoG scaling in *r-new* referents as compared to *r-given* ones.

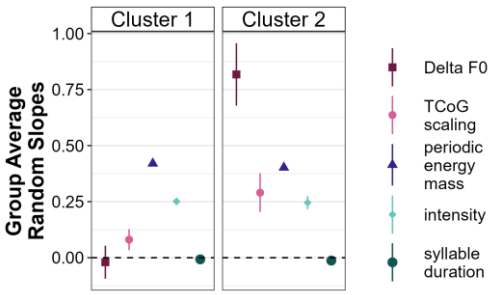

Figure 2: *Results of cluster analysis for referential givenness.*

3.3. Givenness marking and communicative skill

After investigating the nature of individual variability in the prosodic encoding of lexical and referential givenness, we turn to a potential reason for such differences: communicative skill. Participants could score between 0 and 10 points in the questionnaire, where 0 indicates the lowest communicative skill and 10 the highest. Participants scored between 3 and 10 with a mean of 7.3 indicating relatively high communicative skill among most participants. Since the questionnaire we used is part of a screening tool for autism, this distribution may be indicative of a neurotypical sample of participants.

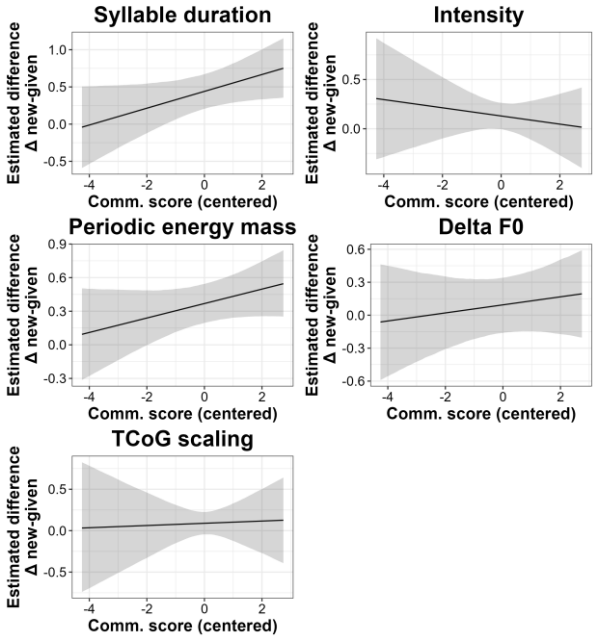

Figure 3: *Estimated differences between lexically new and given referents and 90% credible intervals as a function of communicative score for all parameters.*

In Figures 3 and 4, we present the centered communicative scores, that is, 0 corresponds to the mean across all participants. Figure 3 shows the estimated differences between *l-new* and *l-given* referents and 90% CIs as a function of communicative skill for all five prosodic parameters. In three parameters (i.e., syllable duration, periodic energy mass, and Delta F0), the difference between *l-new* and *l-given* referents clearly increases with higher communicative skill. That is, speakers who self-rated themselves as more communicatively skilled used these

parameters more efficiently in the expected direction to encode a difference in lexical givenness than speakers with lower communicative skill. While for TCoG scaling, there seems to be no moderating effect of communicative skill on the encoding of lexical givenness, the difference in intensity between *l-new* and *l-given* decreases with higher communicative skill.

In Figure 4, the interactions between referential givenness and communicative skill are plotted. Here, we can generally observe the opposite trend to lexical givenness: Speakers identifying as more communicatively skilled make less of a distinction in the expected direction between *r-new* and *r-given* referents in four out of the five prosodic parameters. Only in syllable duration do we observe the expected trend that speakers with higher communicative scores use longer durations in *r-new* than in *r-given* referents.

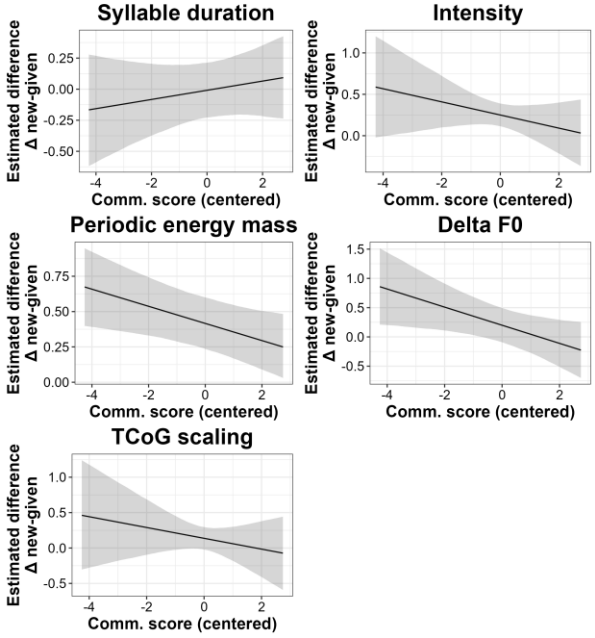

Figure 4: *Estimated differences between referentially new and given referents and 90% credible intervals as a function of communicative score for all parameters.*

4. Discussion and conclusions

We observe different strategies across individuals both in the prosodic marking of lexical and referential givenness. Interestingly, speakers appear to be more similar in the prosodic encoding of the lexical *new-given* contrast. Here, the difference between the two strategies identified in our data lies only in the strength of the encoding, especially via periodic energy mass and duration. In the encoding of referential givenness, by comparison, both clusters use periodic energy mass and intensity to a similar extent, yet one cluster additionally makes use of an increase in F0 for marking *r-new* referents. That is, one cluster encodes the *new-given* contrast more redundantly than the other. The latter cluster does not compensate for the absence of the F0 cue in their productions, leading to weaker referential givenness encoding overall.

One reason for such different individual strategies that has been considered in prominence perception studies is communicative skill. Generally, listeners with higher communicative skill are more adept at perceiving prominence differences than listeners with lower communicative skill [9, 10]. In our production data, we observe contrasting trends for the mediating effect of communicative skill on the encoding of lexical and referential givenness. For lexical givenness, we find the expected trend in that speakers with higher communicative skill more effectively distinguish between *l-new* and *l-given* referents in their productions. However, referential givenness seems to be encoded more clearly by speakers with *lower* communicative skill.

That individual speakers should behave differently in their encoding of lexical givenness on the one hand and referential givenness on the other may be due to the different sources of the two parameters. Referential givenness describes the relation between entities (i.e., creatures, things, states or events) in the situational or textual discourse context, while lexical givenness is determined by the previous occurrence or absence of a word or concept [4]. In fact, our findings are partially in line with [12]'s, who found that speakers with lower communicative skill encoded semantic-pragmatic predictability more strongly than speakers with higher communicative skill. Semantic-pragmatic predictability is comparable to referential givenness in our study. However, Turnbull found no effects of communicative skill regarding second mention reduction, which may be claimed to fall within the scope of lexical givenness.

The interaction of communicative skill and referential givenness is unexpected in the context of previous perception studies that find a higher sensitivity to the mapping of prosody to meaning in communicatively more skilled listeners [9, 10]. In line with these findings, we assumed that highly communicatively skilled speakers should be able to anticipate the higher activation cost demanded of the listener by an *r-new* referent [1] and thus produce a stronger prosodic contrast to clearly mark the pragmatic distinction between *r-new* and *r-given*. It appears that instead, such speakers underestimate this activation cost and generally produce a weaker contrast. A potential explanation is that the communicatively skilled speakers rely more on the morphosyntactic marking of r-givenness, since *r-new* referents are preceded by an indefinite article, while *r-given* referents are preceded by a definite article - making an additional prosodic marking obsolete.

Like [12], we also observe inconsistent mediating effects of communicative skill on different prosodic parameters. It is unclear at this point why higher communicative skill should, for example, have a positive effect on the marking of lexical givenness via duration but not intensity. This remains an issue for further investigation in future studies.

Our findings are limited by the measure of communicative skill we employed, which is very coarse-grained as it relies on the self-assessment of the participants. Other instruments assessing different aspects of communicative skill could yield more precise results that better relate to speech production, e.g., [10] employed a questionnaire to measure emotional and cognitive empathy [22] and found that empathy affects the processing of contrastive focus. Furthermore, we consider relatively few participants (especially compared to perception studies), and most of them had high communicative skill, which limits the generalizability of our findings.

5. Acknowledgements

This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) in the CRC 1252 "Prominence in Language" (Project-ID 281511265).

6. References

[1] W.L. Chafe, *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. University of Chicago Press, 1994.

[2] M. Swerts, E. Krahmer, and C. Avesani, "Prosodic marking of information status in Dutch and Italian: A comparative analysis," *Journal of Phonetics*, vol. 30, no. 4, pp. 629–654, 2002.

[3] C. Féry, and F. Kügler, "Pitch accent scaling on given, new and focused constituents in German," *Journal of Phonetics*, vol. 36, pp. 680-703, 2008.

[4] S. Baumann, and A. Riester, "Coreference, lexical givenness and prosody in German," in J. Hartmann, J. Radó, and S. Winkler (eds.), *Lingua*, vol. 136, pp. 16–37, 2013.

[5] E. Chodroff, and J. Cole, "The phonological and phonetic encoding of information structure in American English nuclear accents," in *Proceedings of the 19th International Congress of Phonetic Sciences*, pp. 1570–1574, 2019.

[6] F. Cangemi, M. Krüger, and M. Grice, "Listener-specific perception of speaker-specific productions in intonation," in S. Fuchs, D. Pape, C. Petrone, and P. Perrier, (eds.), *Individual differences in speech production and perception*, Peter Lang, pp. 123–145, 2015.

[7] M. Grice, S. Ritter, H. Niemann, and T. Roettger, "Integrating the discreteness and continuity of intonational categories," *Journal of Phonetics*, vol. 64, pp. 90-107, 2017.

[8] I.C. Ouyang, and E. Kaiser, "Individual differences in the prosodic encoding of informativity," in S. Fuchs, D. Pape, C. Petrone, and P. Perrier (eds.), *Individual Differences in Speech Production and Perception*, Peter Lang, pp. 147-189, 2015.

[9] J. Bishop, G. Kuo, and B. Kim, "Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: Evidence from Rapid Prosody Transcription," *Journal of Phonetics*, vol. 82, 100977, 2020.

[10] R. Orrico, S. Gryllia, J. Kim, and A. Arvaniti, "The influence of empathy and autistic-like traits in prominence perception," in *Proceedings of the 20th International Congress of Phonetic Sciences*, Prague, Czech Republic, pp. 1280-1284, 2023.

[11] J. Bishop, C. Zhou, K. Antolovic, L. Grebe, K. H. Hwang, G. Imaezue, E. Kistanova, K. E. Lee, K. Paulino, and S. Zhang, "Brief Report: Autistic Traits Predict Spectral Correlates of Vowel Intelligibility for Female Speakers," *Journal of Autism and Developmental Disorders*, vol. 52, no. 5, pp. 2344-2349, 2022.

[12] R. Turnbull, "Listener-oriented phonetic reduction and theory of mind," *Language, Cognition and Neuroscience*, vol. 34, no. 6, pp. 747-768, 2019.

[13] S. Baron-Cohen, R. A. Hoekstra, R. Knickmeyer, and S. Wheelwright, "The Autism Spectrum Quotient (AQ) – Adolescent Version," *Journal of Autism and Developmental Disorders*, vol 35, no. 3, pp. 343-350, 2006.

[14] A. Albert, *A model of sonority based on pitch intelligibility*. Berlin: Language Science Press, 2023.

[15] J. Barnes, N. Veilleux, A. Brugos, and S. Shattuck-Hufnagel, "Tonal Center of Gravity: A global approach to tonal implementation in a level-based intonational phonology," *Laboratory Phonology*, vol. 3, no. 2, pp. 337-383, 2012.

[16] J. Barnes, N. Veilleux, A. Brugos, and S. Shattuck-Hufnagel, "On (and off) ramps in intonational phonology: Rises, falls, and the Tonal Center of Gravity," *Journal of Phonetics*, vol. 85, 101020, 2021.

[17] S. Baumann, and J. Lorenzen, "Boosting or inhibiting - How semantic-pragmatic and syntactic cues affect prosodic prominence relations in German," *PLOS ONE*, to appear.

[18] S. Baumann, and B. Winter, "What makes a word prominent? Predicting untrained German listeners' perceptual judgments," *Journal of Phonetics*, vol 70, pp. 20–38, 2018.

[19] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2022. URL https://www.R-project.org/.

[20] P.-C. Bürkner, "Advanced Bayesian multilevel modeling with the R Package brms," *Journal of Statistical Software*, vol. 80, no. 1, pp. 1-28, 2017.

[21] B. Carpenter, A. Gelman, M. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell, "Stan: a probabilistic programming language," *Journal of Statistical Software*, vol. 76, no. 1, pp. 1–32, 2017.

[22] S. Baron-Cohen, and S. Wheelwright, "The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences," *Journal of Autism and Developmental Disorders*, vol. 34, no. 2, pp. 163–175, 2004.