Interspeech 2025
17-21 August 2025, Rotterdam, The Netherlands

Does effortful speech production indicate communication difficulty caused by noise and hearing aid support?

Lena-Marie Huttner[1,2], *Jeppe H. Christensen*[1], *Gitte Keidser*[1], *Tobias May*[2], *Torsten Dau*[2], *Sergi Rotger-Griful*[1]

[1]Eriksholm Research Centre, Oticon A/S, Snekkersten, Denmark
[2] Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark
leut@eriskholm.com

Abstract

Hearing impairment affects a person's ability to communicate with others. Communication difficulty remains a little understood concept in research. In this study we assess how parameters of speech affected by noise, hearing aid support, and hearing status reflect a participant's experience of communication difficulty. We paired 44 participants into dyads consisting of one normal-hearing (NH) and one hearing-impaired (HI) participant. Participants engaged in task-based conversation in quiet as well as in 70dB background noise. HI participants completed the task both with and without hearing aid. After each conversation participants indicated their communication difficulty on a questionnaire. We then analyze how F1, vocal level, and turn taking variability predict communication difficulty. For HI participants, increased vocal level as well as higher variability in turn taking are associated with greater difficulty. For NH participants, increased F1 is associated with greater difficulty.

Index Terms: Communication difficulty, hearing impairment, conversation in noise

1. Introduction

Hearing impairment affects a person's ability to participate in conversations [1, 2, 3]. Yet, communication difficulty as a measurable concept, remains elusive. Current hearing assessments test a person's hearing ability in an isolated environment which does not resemble the complexities of conversation [4, 5, 6]. While previous studies have identified potential markers in behavior, and speech that are sensitive to changes in noise level and hearing ability, it is unclear how these reflect a person's subjective experience of difficulty in conversation in noise and whether this difficulty can be alleviated by hearing aids (HAs) fitted to hearing-impaired (HI) interlocutors [7]. Previous research has also aimed to understand what participants experience in strenuous conversational settings [8, 9]. However, typically their experience is reported alongside any speech or behavioral measures that are also sensitive to noise and hearing aid amplification and not related back to them.

In this study we observe dyadic task-based conversations between a normal-hearing (NH) and a HI person with acquired hearing loss. We aim to relate participants' subjective experience during conversation in quiet and in noise and with the HI participant unaided and aided to speech measures obtained during the dyadic conversations. Previous research has shown that speech and conversation behavior changes for both NH and HI participants when conversing in different levels of background noise [10, 11].

Firstly, in noise, speakers tend to raise their vocal level [7]. [9] Further found that HI participants' vocal level was modified by the prescribed amplification: They tended to speak louder when unaided, and less loud when aided. Vocal intensity is therefore a reliable marker associated with speech in challenging conditions. Secondly when speaking in noisy settings speakers tend to hyperarticulate. While this increase in effort is indicated by several acoustic parameters[11, 12], increases in F1 have repeatedly been associated with increasing noise levels, and are generally less prone to gender-based differences than F0 [3]. [12] find that F1 is especially sensitive to increases in noise level when the background noise is speech shaped. They ([12]) further find that this adaptation increases the intelligibility of the produced signal. Thirdly, higher levels of linguistic processing are affected by challenging speaking conditions. Turn taking behavior has repeatedly been shown to be sensitive to background noise and hearing aid amplification: In particular [7] and [9] show that variability in floor transfer offset (FTO) is affected by background noise and hearing ability. This measure is reflective of a participant's inability to predict the end of a turn and plan their next utterance [9]. When conversing in noise variability in FTO increased for all participants relative to conversation in quiet, but significantly more for HI than NH participants. While these speech parameters, vocal level, F1 and FTO variability are associated with more challenging test conditions and are modifiable by HA amplification, it is unclear whether these adaptations are indicative of a participant's experience of communication difficulty.

Communication difficulty is a multifaceted concept [8, 13]. To date there are no standardized measurement instruments to assess communication difficulty. But subjective reports of participant experience have been examined in several studies. For example, in [7] HI and NH participants were asked to rate their communicative effort after conversing at different levels of background noise. While effort was negatively affected by noise it was not predicted by increased vocal intensity. [7] assessed effort using a single question, acknowledging that this may not suffice to capture this complex phenomenon [8, 13]. In [10] a 10-item questionnaire and subsequent principal component analysis was used to assess factors contributing to difficulty, relevance, and engagement in a task-based conversation. Here, perceived difficulty increases with

10.21437/Interspeech.2025-1535

increasing noise level. Both [8] and [10] find that difficulty ratings are not affected by a participant's hearing status.

[14] employed a questionnaire – a subset of the NASA Task Load Index (TLX) [15] to gage participants' frustration, effort and success in dyadic conversation. The NASA TLX is a validated questionnaire to assess task demand that may be an apt tool to better understand communication difficulty in task-based dialogue. [13] employed a qualitative approach to further define communication success and found that engagement and effort were indicative of success for HI participants in dialogue. [8] further found that communication success was affected by noise and sensitive to HA support.

The goal of this study is to investigate how vocal level, F1, and FTO variability (as indicated by the inter quartile range) in conversation predict a participant's subjective experience of communication difficulty as measured by a questionnaire at different noise levels and with and without hearing aid support provided to the HI interlocutor. As reported by [3] and [7] we expect both HI and NH participants to report greater difficulty as background noise increases. We hypothesize that difficulty can be alleviated by hearing aid support for HI participants. We further investigate which speech markers are associated with subjective ratings of difficulty for both HI and NH participants.

2. Methods

2.1. Participants

We recruited forty-four older native Danish-speaking individuals. Participants were grouped into twenty-two dyads consisting of one NH and one HI participant each. The NH participants, aged between 55 and 75 years (mean = 65.8 years, SD = 6.6 years, 11 female, 11 male), demonstrated age-related normal hearing thresholds as defined by ISO-7029. Their pure tone average (PTA) thresholds ranged from -1 to 33.5 dB HL as measured across 250, 500, 1000, and 2000 Hz and across left and right ears. The HI participants were aged between 51 and 80 years (mean = 72.5 years, SD = 8.6 years, 11 female, 11 male), and had symmetrical PTA thresholds between 31.3 to 56.3 dB HL. Participants were grouped into mixed gender dyads, consisting of one male and one female participant. All HI participants were active hearing aid (HA) users with a minimum of 6 months experience. Participants exhibited no cognitive impairment, tinnitus, or fluctuating hearing loss. The study received ethics approval from the Science-Ethics committee of the Capital Region of Denmark (reference number H-16036391).

2.2. Test setup

Each pair engaged in a face-to-face dyadic conversation elicited by working together to identify up to 12 differences between two Diapix pictures (near-identical pictures) sourced from the original Diapix corpus [16] translated to Danish where applicable. The laboratory was equipped with eight equally spaced loudspeakers in a horizontal ring, Participants were seated facing each other on opposite sides of a table positioned in the center of the loudspeaker array. They were equipped with head-worn directional microphones (DPA 4488, Germany) to record their speech. The communication task was carried out in quiet (N0 condition) and with 70dB 10-talker babble noise played back through the loudspeaker array (N70 condition), and HA support (unaided and aided condition) provided to HI participants. In the aided condition, HI participants were fitted with identical HAs using the NAL-NL2 prescription [17] The fit to target was verified through real-ear measurements. Consequently, pairs completed four communication trials in randomized order: both interlocutors unaided at noise levels of N0 and N70, and with the HI interlocutor aided at noise levels of N0, and N70. To assess the participants' subjective experience of communication difficulty, we provided them with a questionnaire after each conversation in which they indicated their frustration, effort, success, stress, and engagement. We here adapted the questionnaire employed by [14] and expanded it by the items of engagement [13] and stress.

2.3. Data processing

To extract the mean intensity during the conversation, we employed the Voice Activity Detection (VAD) algorithm described by [18]. To correct for any background noise picked up by the microphones, all intensity measures were corrected by 7 dB in the noisy condition in line with the method described by [19]. To obtain the inter quartile range (iqr) of FTO, we classified overlaps and gaps between turns using an algorithm designed by [18]. We then subtracted the duration of overlaps from the duration of gaps. F1 was obtained through the linear predictive coding technique [20]. All data processing was performed in MATLAB.

2.4. Statistical analysis

All analyses were carried out in R [21]. We assume that the five questions from the questionnaire capture different aspects of the same latent concept: communication difficulty. To analyze communication difficulty we reduce the items to a single factor using a principal component analysis (PCA) with varimax rotation [10]. In this procedure we extract one factor, which explains 71% of the variance in the data. The factor loadings of the individual items are shown below in table 1. The extracted principal component is henceforth referred to as communication difficulty.

Table 1: Factor loadings of the individual questionnaire items

Variable	Factor loading
Engagement	0.797
Frustration	0.895
Stress	0.873
Success	-0.779
Effort	0.865

We then build linear mixed effects models (lmems) using the LmerTest package [22]. We first investigate whether communication difficulty varies by noise level, hearing aid support, and hearing status. In a second step of analysis we then examine whether communication difficulty is predicted by vocal level, F1, or iqrFTO. In this analysis all speech measures were scaled using the *scale* function in R. To reduce the parameters and improve model fit we follow the recommendations of [23] to select the best fit model via parametric model comparison based on the corrected Akaike Information Criterion (AICc) [24] To do so we first fit the maximal model including all three speech variables as fixed factors and then iterate through all possible factor combinations using the MuMIN package [25]. We then select the model that includes the most fixed factors within delta AICc < 2 of the lowest AICc model.

3. Results

In a first step we analyze whether noise, hearing status and hearing aid setting influenced the subjective difficulty scores. Figure 1 shows the difficulty scores per noise level and participant group in each aid condition. In this figure, communication difficulty (y-axis) refers to the extracted principal component. A higher value on this scale indicates more communication difficulty.

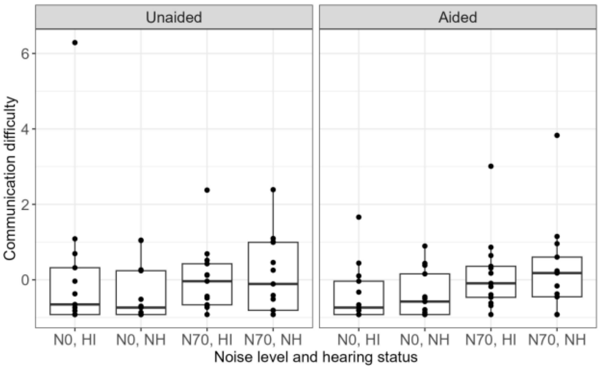

Figure1: *difference in communication difficulty ratings as indicated by the PCA by noise level between participant groups for each aid level*

In a first model investigating the effect of condition, aid and hearing status we predict subjective difficulty with hearing (NH/HI), noise level (N0/N70) and HA support (Aided/Unaided) up to the second order interaction and included participant as a random intercept (formula: difficulty ~ hearing * noise level * HA support + (1|participant)). The explanatory power of the model is moderate with a conditional R^2 of 0.65 and a marginal R^2 of 0.05. Main effects were calculated using ANOVA. The results reveal that noise level had a significant effect on the subjective difficulty rating (F =8.3640, p = 0.005). Neither hearing status nor HA support support or their interactions had a statistically significant effect. This indicates that subjective difficulty is affected by noise, whereas HA support has no effect. There is no effect of hearing status on subjective difficulty. While this indicates that both participant groups had similar experiences during the conversation, HI and NH participants may have employed different compensatory behaviors indicative of communication difficulty. In the second step of analysis we therefore build lmems predicting communication difficulty with F1, iqrFTO, and vocal level for each participant group.

HI model

We first analyze the data for the HI participants. We built a maximal model predicting difficulty with F1, iqrFTO, and vocal level as fixed factors and participants as random intercept (difficulty~ F1 + iqrFTO + vocal level + (1|participant)). We then ran the model selection function. The model with the lowest AICc included iqrFTO and vocal level as fixed factors (difficulty~ iqrFTO + vocal level + (1|participant). In this model the conditional R^2 was 0.80 indicating that fixed and random effects explain 80% of the variance in the data. The fixed effects alone account for 11% of the variance with a marginal R^2 of 0.11. The normality of residuals was inspected via Q-Q-plot. IqrFTO is a significant predictor of communication difficulty, with a positive effect, (β=0.346, t(40.12)=3.818t(40.12) = 3., p=0.0005). Vocal level was also a statistically significant and positive predictor of communication difficulty (β=0. 223, t(39.09)=2.253, p=0.03). Figures 2 and 3 below illustrate the relationship between the subjective difficulty score and vocal level (figure 2 - left) and iqrFTO (figure 2 - right).

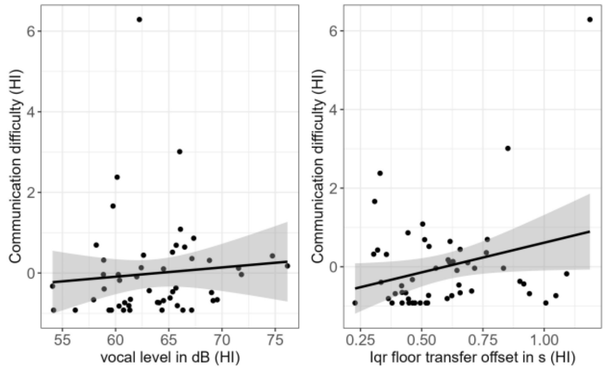

Figure 2: *Relationship between communication difficulty and vocal level in dB (left) and Iqr FTO in seconds (right) for the hearing-impaired participants.*

NH model

We then analyze the data of the NH participant group in the same manner. Employing the same procedure, we built a maximal model predicting difficulty with iqrFTO, F1, and vocal level while including participant as a random intercept. The model with the lowest AICc included only F1 as a fixed factor and a random intercept for participant (difficulty ~ F1 + (1|participant)). The assumption of normality of residuals was inspected via Q-Q-plot. The model accounts for 55% of the variance in the data (conditional $R^{2=}$ 0.55), with the fixed effects alone explaining 10 % (marginal R^2 =0.10). F1 was a statistically significant and positive predictor of communication difficulty (β=0.320, t(42.28) =3.043, p=0.004). below, in figure 3 we depict the relationship between F1 and subjective difficulty for the NH group.

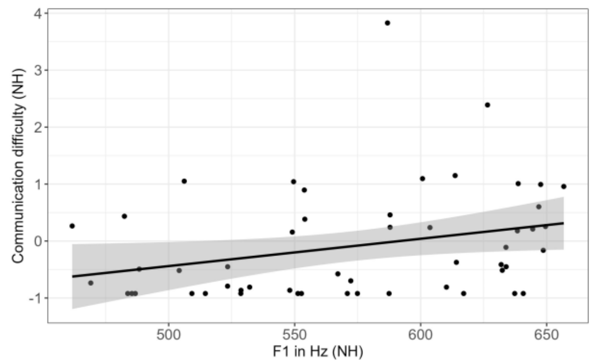

Figure 3: *Relationship between communication difficulty and F1 in Hz for NH participants.*

4. Discussion

We find that communication difficulty as indicated by subjective responses to a questionnaire, is affected by noise while neither hearing aid support nor hearing status have an effect on the subjective difficulty score (see figure 1). This is in line with findings by [10], suggesting that interlocutors engaging in a dyadic conversation, despite having different hearing abilities, report similar levels of communication difficulty. Although a slight tendency of difficulty improving with HA support for HI participants can be seen in figure 1, this tendency is not statistically significant.

We further investigated whether vocal level, F1, and FTO variability are predictors of communication difficulty in noise. The results show that the speech parameters associated with difficulty differ by participant group. Much like noise leads to different adaptive behaviors for different participant groups, as found by [7], different speech and turn taking markers are associated with difficulty for HI and NH participants.

The results indicate that for HI participants greater difficulty is associated with increased variability in FTO (see figure 2). The effect of turn-taking behavior on perceived communication difficulty when HI participants are challenged by relatively high levels of noise is supported by previous findings that turn-taking duration and variability increase in HI participants when communicating in increasing levels of noise [7, 9]. Increased background noise may make it difficult for HI participants to predict the end of a turn and prepare an utterance thus inhibiting conversational flow – a concept further associated with difficulty as indicated by [8] and [13].

Increased vocal level predicts communication difficulty for the HI group, but this effect is weaker than the effect of iqrFTO. For NH participants, an increase in F1 is associated with difficulty (see figure 3). Previous research has also found that both vocal level and F1 tend to increase in noise [3]. This is further exhibited in the Lombard effect – a strong involuntary speech response to noise [11] generally associated with own-speech perception. The results reported herein indicate that this compensatory mechanism is predictive of communication difficulty. The two participant groups appear to employ different compensatory strategies that are associated with the Lombard effect. While difficulty for HI participants is associated with louder speech, as reported in [3] NH participants appear to raise F1 suggesting hyperarticulation. An increase in F1 specifically has been associated with greater intelligibility in noise [12]. Moreover, [11] argue that Lombard speech could have a communicative effect, signaling discomfort to the interlocutor. An adaptation of F1 in NH participants could be a response to the HI participant who may not only speak louder to hear themselves better but also signal difficulty in distinguishing speech sounds in noise. This may suggest an asymmetry in the function of markers indicating difficulty between the two participant groups. While such an asymmetry has been suggested by [3] and [11], it may be a fruitful avenue for future research to investigate the function and perception of markers indicating difficulty.

Communication difficulty remains hard to operationalize due to the lack of scientific consensus. We agree with [7] and [13] who emphasize that single item questionnaires may not capture the full picture. While we employed an established questionnaire expanded by items frequently associated with communication difficulty [14, 15, 13] we still may not have aptly captured a participant's experience of difficulty. We also cannot fully infer the source of communication difficulty. While communication difficulty varies by noise level, we cannot comment on any other internal (e.g. cognitive factors, general discomfort) or external factors (e.g. the interlocutor's difficulty) that may have contributed to the participants' experience. We advocate for further qualitative and quantitative exploration of a HI person's experience to better inform empirical approaches.

Furthermore, we here limited our analysis to three parameters of conversational speech which have repeatedly and reliably been associated with effortful conversational speech in challenging settings [3, 11, 12, 18]. We, however, recognize that participants might have employed compensatory mechanisms that we did not analyze. This study might be underpowered to reliably detect the effects of factors that are more prone to gender or individual differences like F0. The findings in this study are limited by constraints in our sample. We have limited data in this experiment with 44 total observations, 22 per participant group. However, the sample size is comparable with other studies in the field, and subject to design constraints regarding recruiting and scheduling. In this experiment, participants only held one conversation in each condition. Future studies may be advised to include more repetitions to increase statistical power.

The cohort of NH exhibited very mild hearing loss that nevertheless were accepted as falling within the normal range for their age. While there was a significant difference between the audiograms of both participant groups we cannot completely rule out that very mild hearing loss in the NH group may have influenced the results and the participants' behavior. Other findings may be obtained when studying dyadic conversations among other configurations of interlocutors, particularly if pairing an older HI and a young NH person, who beside different ages and hearing ability may also display differences in cognitive function and language skills.

Furthermore, participants in this study engaged in task-based dialogue. It is likely that communication difficulty in a task-based conversation is evaluated and experienced differently than communication difficulty in free conversation [26, 27] which may more closely resemble real-life conversations. We therefore recommend that future studies include a free conversation task to corroborate the results.

In this study we related the participants' communication difficulty to their own behavioral data. In line with [10] and [11] we assume that a participant will voluntarily or involuntarily adapt their own behavior to both alleviate discomfort and signal difficulty to their interlocutor. We recognize that conversation is a complex, synergistic behavior in which interlocutors adapt not only to a situation but also dynamically to one another [27]. Both a participant's behavior as well as their experienced communication difficulty may have been affected by their interlocutor's behavior. While an analysis to this end is beyond the scope of this study we strongly encourage future research to examine synergistic effects of communication difficulty. Nevertheless the results suggest that objective speech measures may predict whether a person is experiencing communication difficulty. The speech measures indicating difficulty may differ between NH and HI persons.

5. Conclusion

Hearing impaired people often seek hearing care because of experienced communication difficulty, particularly in noisy environments. Both clinically and in research communication difficulty remains ill-defined. We here evaluate communication difficulty as a latent variable established from questions on success, effort, engagement, stress and frustration. We show that objective speech measures, like vocal level, F1, and FTO, may be indicative of an older person's subjective experience of communication difficulty when conversing with peers in a noisy environment. The latent variable was not sensitive to hearing ability or hearing aid support.

6. Acknowledgements

This research was supported by the Innovation Foundation Denmark (1045-00007B, 3195-00018B) and the William Demant Foundation. The grant was awarded to Sergi Rotger-Griful.

7. References

[1] N. Williams, D. M. Guthrie, J. G. S. Davidson, K. Fisher, and L. E. Griffith, "A Deterioration in Hearing Is Associated With Functional and Cognitive Impairments, Difficulty With Communication, and Greater Health Instability," *Journal of Applied Gerontology*, vol. 39, no. 2, pp. 159–171, Feb. 2020, doi: 10.1177/0733464818755312.

[2] S. Aliakbaryhosseinabadi, G. Keidser, T. May, T. Dau, M.A. Skoglund, and S. Rotger-Griful, "Measuring communication difficulty with eye-gaze behavior when speaking and listening. ," in *49. Jahrestagung für Akustik.*, 2023.

[3] T. Beechey, J. M. Buchholz, and G. Keidser, "Measuring communication difficulty through effortful speech production during conversation," *Speech Commun*, vol. 100, pp. 18–29, Jun. 2018, doi: 10.1016/j.specom.2018.04.007.

[4] Y.-H. Wu, E. Stangl, O. Chipara, S. S. Hasan, S. DeVries, and J. Oleson, "Efficacy and Effectiveness of Advanced Hearing Aid Directional and Noise Reduction Technologies for Older Adults With Mild to Moderate Hearing Loss," *Ear Hear*, vol. 40, no. 4, pp. 805–822, Jul. 2019, doi: 10.1097/AUD.0000000000000672.

[5] T. Beechey, "Is speech intelligibility what speech intelligibility tests test?," *J Acoust Soc Am*, vol. 152, no. 3, pp. 1573–1585, Sep. 2022, doi: 10.1121/10.0013896.

[6] S. Carlile and G. Keidser, "Conversational Interaction Is the Brain in Action: Implications for the Evaluation of Hearing and Hearing Interventions," *Ear Hear*, vol. 41, no. Supplement 1, pp. 56S-67S, Nov. 2020, doi: 10.1097/AUD.0000000000000939.

[7] E. B. Petersen, E. N. MacDonald, and A. Josefine Munch Sørensen, "The Effects of Hearing-Aid Amplification and Noise on Conversational Dynamics Between Normal-Hearing and Hearing-Impaired Talkers," *Trends Hear*, vol. 26, Jan. 2022, doi: 10.1177/23312165221103340.

[8] R. Nicoras *et al.*, "Small Group Conversations: Communication Behavior and Success," Turin, Sep. 2023.

[9] E. B. Petersen and D. Parker, "Speak up: How hearing loss and the lack of hearing aids affect conversations in quiet," *Journal of Speech, Language, and Hearing Research*, vol. 67, no. 8, pp. 2782–2793, Aug. 2024, doi: 10.1044/2024_JSLHR-23-00667.

[10] T. Beechey, J. M. Buchholz, and G. Keidser, "Eliciting Naturalistic Conversations: A Method for Assessing Communication Ability, Subjective Experience, and the Impacts of Noise and Hearing Impairment," *Journal of Speech, Language, and Hearing Research*, vol. 62, no. 2, pp. 470–484, Feb. 2019, doi: 10.1044/2018_JSLHR-H-18-0107.

[11] T. Beechey, J. M. Buchholz, and G. Keidser, "Hearing Impairment Increases Communication Effort During Conversations in Noise," *Journal of Speech, Language, and Hearing Research*, vol. 63, no. 1, pp. 305–320, Jan. 2020, doi: 10.1044/2019_JSLHR-19-00201.

[12] Y. Lu and M. Cooke, "Speech production modifications produced by competing talkers, babble, and stationary noise," *J Acoust Soc Am*, vol. 124, no. 5, pp. 3261–3275, Nov. 2008, doi: 10.1121/1.2990705.

[13] R. Nicoras, S. Gotowiec, L. V Hadley, K. Smeds, and G. Naylor, "Conversation success in one-to-one and group conversation: a group concept mapping study of adults with normal and impaired hearing," *Int J Audiol*, vol. 62, no. 9, pp. 868–876, Sep. 2023, doi: 10.1080/14992027.2022.2095538.

[14] S. Aliakbaryhosseinabadi, G. Keidser, T. May, T. Dau, D. Wendt, and S. Rotger-Griful, "The Effects of Noise and Simulated Conductive Hearing Loss on Physiological Response Measures During Interactive Conversations," *Journal of Speech, Language, and Hearing Research*, vol. 66, no. 10, pp. 4009–4024, Oct. 2023, doi: 10.1044/2023_JSLHR-23-00063.

[15] S. G. Hart, "Nasa-Task Load Index (NASA-TLX); 20 Years Later," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, pp. 904–908, Oct. 2006, doi: 10.1177/154193120605000909.

[16] R. Baker and V. Hazan, "DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs," *Behav Res Methods*, vol. 43, no. 3, pp. 761–770, Sep. 2011, doi: 10.3758/s13428-011-0075-y.

[17] G. Keidser, H. Dillon, L. Carter, and A. O'Brien, "NAL-NL2 Empirical Adjustments," *Trends Amplif*, vol. 16, no. 4, pp. 211–223, Dec. 2012, doi: 10.1177/1084713812468511.

[18] A. J. M. Sørensen, M. Fereczkowski, and E. N. MacDonald, "Effects of Noise and Second Language on Conversational Dynamics in Task Dialogue," *Trends Hear*, vol. 25, Jan. 2021, doi: 10.1177/23312165211024482.

[19] L. V. Hadley, W. O. Brimijoin, and W. M. Whitmer, "Speech, movement, and gaze behaviours during dyadic conversation in noise," *Sci Rep*, vol. 9, no. 1, p. 10451, Jul. 2019, doi: 10.1038/s41598-019-46416-0.

[20] R. C. Snell and F. Milinazzo, "Formant location from LPC analysis data," *IEEE Transactions on Speech and Audio Processing*, vol. 1, no. 2, pp. 129–134, Apr. 1993, doi: 10.1109/89.222882.

[21] Posit team, "RStudio: Integrated Development for R. RStudio," 2024, *Posit Software, PBC, Boston, MA*.

[22] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," *J Stat Softw*, vol. 82, no. 13, 2017, doi: 10.18637/jss.v082.i13.

[23] H. Matuschek, R. Kliegl, S. Vasishth, H. Baayen, and D. Bates, "Balancing Type I error and power in linear mixed models," *J Mem Lang*, vol. 94, pp. 305–315, Jun. 2017, doi: 10.1016/j.jml.2017.01.001.

[24] C. M. Hurvich and C.-L. Tsai, "Regression and time series model selection in small samples," *Biometrika*, vol. 76, no. 2, pp. 297–307, 1989, doi: 10.1093/biomet/76.2.297.

[25] "MuMIn: Multi-Model Inference," May 28, 2010. doi: 10.32614/CRAN.package.MuMIn.

[26] C. J. Wynn, T. S. Barrett, and S. A. Borrie, "Conversational Speech Behaviors Are Context Dependent," *Journal of Speech, Language, and Hearing Research*, vol. 67, no. 5, pp. 1360–1369, May 2024, doi: 10.1044/2024_JSLHR-23-00622.

[27] R. Fusaroli, J. Rączaszek-Leonardi, and K. Tylén, "Dialog as interpersonal synergy," *New Ideas Psychol*, vol. 32, pp. 147–157, Jan. 2014, doi: 10.1016/j.newideapsych.2013.03.005.