DOI: 10.61782/fa.2023.0216

EFFECT OF NOISE, DYSPHONIC VOICE AND INDIVIDUAL DIFFERENCES ON SPEECH INTELLIGIBILITY AND COMPREHENSION

Silvia Murgia[1*] **Bisma N. Choudhry**[1] **Mary M. Flaherty**[1] **Pasquale Bottalico**[1]

[1] Department of Speech and Hearing Science, University of Illinois at Urbana-Champaign, United States

ABSTRACT

Noisy classrooms and the poor voice quality of the teacher decrease children's speech understanding and comprehension. Emerging evidence demonstrates that cognitive factors such as working memory (WM), attention, and inhibitory control (IC) also mediate children's ability to recognize speech in noise and dysphonic voices. Speech recognition and comprehension tests were performed in a sound-proof booth with 15 normal-hearing elementary students. The speech material was recorded with a normal and mimicked dysphonic voice. Babble noise was added to obtain 2 signal-to-noise (SNR) ratios at 12 and 0 dB. Recognition, comprehension and subjective listening effort scores were collected and children's WM and IC were evaluated to explain individual differences. Results showed a statistically significant decrease in recognition performance when SNR was decreasing and in comprehension performance in the presence of the dysphonic voice. Subjective listening effort increased for both lower SNR and in the presence of dysphonic voice. Finally, greater selective attention was associated with better recognition and comprehension performance while greater working memory capacity was associated with decreased listening effort.

Keywords: *speech recognition, speech comprehension, dysphonic voice, background noise, children.*

**Corresponding author*: *smurgia2@illinois.edu.*

1. INTRODUCTION

Noisy classrooms with poor acoustics are common across the U.S. and these unfavorable acoustic conditions result in difficulty in understanding speech. [1–4] Additionally, up to 58% of teachers experience voice disorders during their careers. [5] Children have greater difficulty than adults when understanding speech in the presence of competing sounds, [6,7] and this can have negative impacts on learning and academic achievement. [8–10] Emerging evidence demonstrates that cognitive factors such as working memory, attention, and inhibitory control also mediate children's ability to recognize speech in noise [11,12] and understand dysphonic voices. [13] Current evaluations of speech recognition ability are limited to clinic or laboratory settings and do not take into account the classroom environment. The validity of traditional speech recognition tests for detecting hearing loss and predicting classroom performance has not been established, and speech-in-noise assessments are not widely used in pediatric clinical settings. Additionally, the impact of classroom noise on higher-level cognitive processes such as listening comprehension, which is critical for academic success and literacy development, needs to be examined. Listening comprehension requires the listener to understand and respond to the content of a sentence or passage. This represents higher-level linguistic processing and interpretation of the syntactic and semantic properties and lexical content of the sentence. This requires the parsing of words using syntax, integrating new information with lexical knowledge, and drawing inferences based on semantic information in long-term memory. [14] Listening comprehension more closely approximates the functional listening abilities that are required to achieve academic success and develop age-appropriate literacy skills. [15,16] Evidence is limited but

suggests that noise can have a negative impact on children's listening comprehension, at least in a laboratory setting. [15–17] Given that the acoustic environment not only impacts the signal quality but also increases cognitive processing demands related to attention, inhibitory control, and working memory, the impact of classroom noise has consequences beyond decreased recognition scores, such as increased listening effort and decreased memory capacity. [10, 13, 17, 18] It is, therefore, critical to examine how these factors interact and build the foundation to predict performance in different classroom conditions and how they relate to higher-level cognitive processes.
The following research questions will be addressed:

1. Do background noise and a dysphonic voice have the same effect on speech recognition and speech comprehension?
2. Do background noise and a dysphonic voice have a more detrimental effect on listening effort compared to accuracy for word recognition and listening comprehension?
3. Do cognitive abilities play a role in speech recognition, speech comprehension, and listening effort in noise when the speech has a normal quality?
4. Do cognitive abilities play a role in speech recognition, speech comprehension, and listening effort in noise when the speech is dysphonic?

2. METHODS

2.1 Participants

The participants consisted of 14 students between 8 and 12 years old (mean=10.4 SD= 1.7). The students were equally distributed between males (7) and females (7). All children received a hearing screening from an audiologist, and all showed a normal hearing status, with thresholds of $\leq$ 20 dB HL for octave frequencies between 250 and 8000 Hz. [19] All participants and their parents signed informed consent for their participation in the study, which was approved by the Institutional Review Board of the University of Illinois Urbana-Champaign under Protocol No. 19120.

2.2 Procedure

The experiment took place in a soundproof booth. The children performed the test using a computer and listening to the speech material through headphones. Speech recognition and comprehension were measured using a closed-set, picture-pointing procedure. For speech recognition, the target stimuli were selected from the Word Intelligibility by Picture Identification (WIPI) [20] word list. This test is composed of monosyllabic words which are appropriate for school-age children. The target stimuli for speech comprehension consisted of the Test for Reception of Grammar Version 2 (TROG-2). [21] The TROG-2 is suitable for children aged 4 to 18+ and assesses verbal language comprehension through 20 blocks each with a different grammatical contrast with increasing linguistic complexity. After each trial children rated the subjective effort experienced on a 5 points visual scale from 0 (very easy) to 4 (very difficult).

2.3 Speech material and conditions

The speech material was recorded by a female professional voice user and native speaker of American English. She pronounced the sentences with normal voice quality and with a simulated dysphonic voice after being instructed by a speech-language pathologist to mimic a breathy voice. Classroom noise was added to simulate the noise generated in a real classroom due to the kids talking and moving around. Two SNRs were selected (12 dB(A) and 0 dB(A)) to simulate a typical classroom environment. 4 listening conditions were generated consisting of 2 voice qualities (normal vs. dysphonic) and 2 SNRs.

2.4 Cognitive assessment

Cognitive abilities were assessed with two tests from the NIH Toolbox Cognition Battery performed on an iPad: (1) NIH Toolbox List Sorting Working Memory Test, and (2) NIH Toolbox Flanker Inhibitory Control and Attention Test. The list Sorting Working Memory Test assesses children's working memory (WM). Stimuli were presented visually and auditory. The children were asked to repeat the words in size order from the smallest to the largest. The Flanker Inhibitory Control and Attention Test was used to assess children's attention and inhibitory control. The participants were asked to focus on a target stimulus inhibiting attention to the other distracting stimuli. The children were presented with a row of arrows and they were asked to select the direction in which the middle arrow was pointing.

2.5 Statistical analysis

Generalized linear models (GLM) were applied for the statistical analysis of recognition score, comprehension score, and subjective listening effort (LE) using the software R3.6.0 and the lme4 (version 1.1–10) package. [22]

The LE response variable (5-point scale from 0 to 4) was recoded dividing each value by 4 to restrain the range between 0 and 1. The independent variables included in each model were (1) voice quality, (2) SNR, (3) WM, and (4) IC. To analyze WM ad IC the uncorrected standard scores have been used. The GLM outputs include the estimates of the fixed effects of the coefficients, the standard error associated with the estimate, the test statistic, z, and the p-value.

3. RESULTS AND DISCUSSION

3.1 Speech recognition and comprehension accuracy

The statistical analysis of speech recognition indicated a main effect of the noise (Estimate=-1.97, $p<0.001$) where decreasing SNR resulted in a significant decrease in accuracy regardless of voice quality. The relationship between speech recognition and SNR, grouped by voice quality, is shown in Fig. 1. Additionally, the model revealed an effect of inhibitory control and attention capacity of children (Estimate=0.08, $p=0.02$). Children who had a reduced ability to selectively focus on the target speech and filter out noise distractions performed worse in speech recognition tasks. The statistical analysis of speech comprehension indicated an effect of the dysphonic voice (Estimate=-0.71, $p<0.001$). Children's comprehension accuracy decreased when the sentences were presented with a dysphonic voice. The relationship between Speech comprehension and SNR, grouped by voice quality, is shown in Fig. 2. As found for the speech recognition task, inhibitory control and attention capacity of children showed a significant relationship with the speech comprehension scores (Estimate=0.06, $p=0.007$). Children who exhibit higher levels of selective attention were capable of better concentration which resulted in higher comprehension of sentences as opposed to their peers with lower scores. The results of the interaction between the accuracy scores and the cognitive abilities of the children highlighted the critical role of selective attention in both recognition and comprehension tasks where high levels of background noise are present.

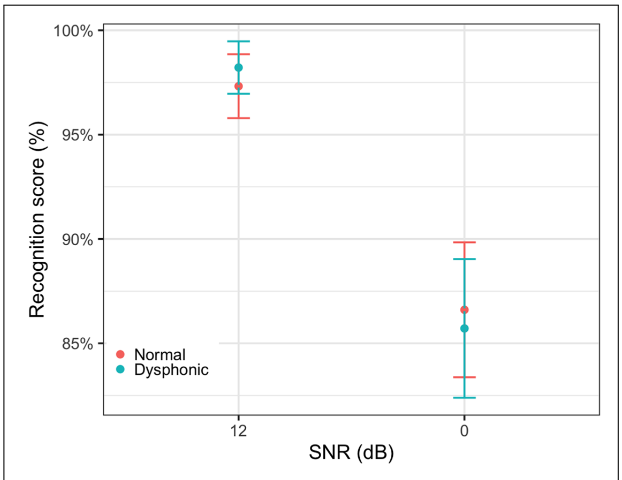

Figure 1. Mean speech recognition scores in % across subjects. The error bars indicate ±standard error.

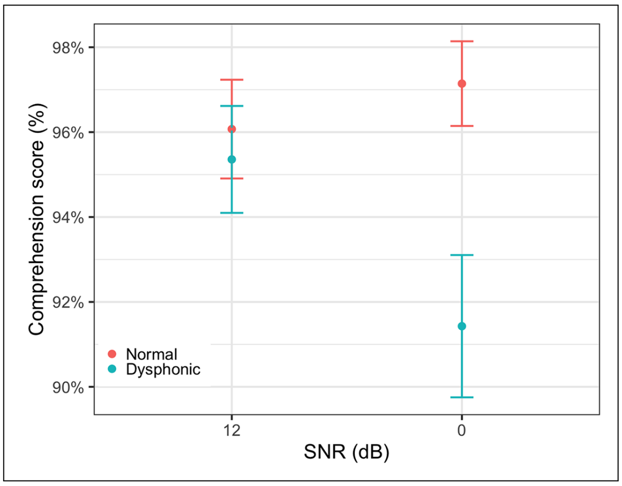

Figure 2. Mean speech comprehension scores in % across subjects. The error bars indicate ±standard error.

3.2 Subjective listening effort

The statistical analysis of the subjective listening effort for the speech recognition task demonstrated a main effect of both SNR (Estimate=0.58, $p<0.001$) and voice quality (Estimate=0.80, $p<0.001$). In the presence of the highest level of noise and dysphonic speech, the children perceived a greater effort in listening to the words. The relationship between LE and SNR, grouped by voice quality, is shown in Fig. 3. The subjective rating of listening effort has been shown to be more sensitive than the test accuracy as even when children understood the same number of words in the presence of dysphonic speech compared to normal speech, they rated the disphonic voice quality as more difficult to listen to. The model also revealed a main effect of working memory on listening effort (Estimate=-1.12, $p<0.001$). Children exhibiting higher working memory capacity were observed to experience less effort in listening to speech in the presence of a higher noise level and the dysphonic voice as compared to their counterparts with lower working memory capacity.

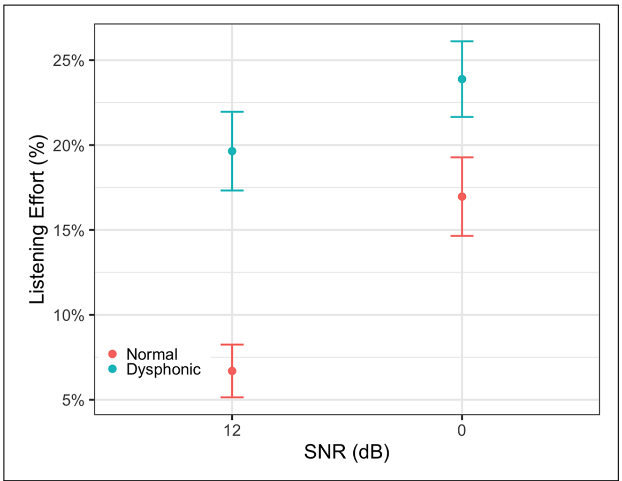

Figure 3. Mean scores of subjective listening effort in % for the speech recognition task. The error bars indicate ±standard error.

The statistical analysis of the subjective listening effort for the speech comprehension task showed a main effect of both SNR (Estimate=0.52, $p<0.001$) and voice quality (Estimate=0.74, $p<0.001$). Increased levels of noise and the presence of dysphonic speech led to a significant increase in perceived listening effort in the speech comprehension task. The relationship between LE for the speech comprehension task and SNR, grouped by voice quality, is shown in Fig. 4. Listening effort was found again to be a more sensitive metric, as the accuracy of the speech comprehension task was only impacted by the presence of a dysphonic voice. However, the children perceived a greater listening effort even when the noise levels were higher. Regarding cognitive abilities, the model showed a main effect of both working memory (Estimate=-0.06, $p<0.001$) and inhibitory control (Estimate=0.07, $p<0.001$). With an increase in their working memory capacity, the children needed to expend less effort while listening since they were able to retain the information more easily and manipulate it so as to understand the sentence described. On the other hand, the increase in selective attention capacity resulted in greater listening effort. This could be attributed to the children's greater awareness of the difficulty of the task and their stronger willingness to concentrate under challenging conditions, which ultimately led to a greater sense of perceived effort. A noteworthy limitation of the study is that despite the presence of poorer SNR and the dysphonic voice, the speech recognition and comprehension scores remained quite high. The high levels of performance achieved in this context may have impacted the sensitivity of accuracy measurements compared to the listening effort. Another limitation to consider is the small sample size utilized in the study. Consequently, it was not feasible to examine the impact of age, which will be addressed in future analyses with a larger sample size.

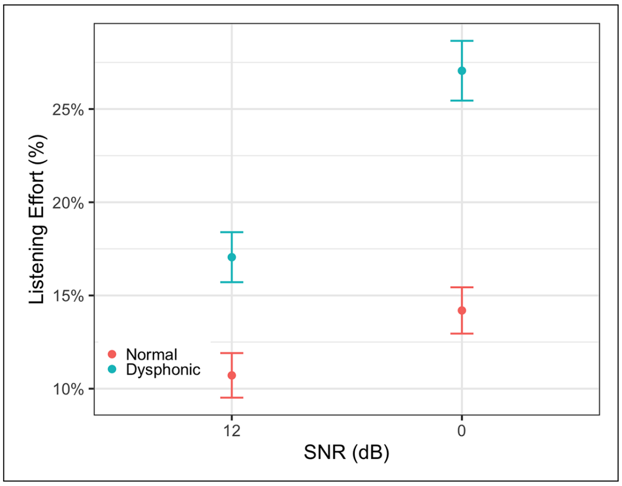

Figure 4. Mean scores of subjective listening effort in % for the speech comprehension task. The error bars indicate ±standard error.

4. CONCLUSION

In conclusion, speech recognition was affected by the SNR while speech comprehension was impacted by the dysphonic voice. In both tasks, increased accuracy was associated with better children's selective attention. The subjective listening effort was shown to be a more sensitive measure compared to the accuracy score since it was impacted by SNR and the voice quality for both recognition and comprehension tasks. Additionally, children with higher working memory scores reported lower listening effort. However, for the speech comprehension task, children with higher selective attention reported greater listening effort compared to children with poorer selective attention.

5. ACKNOWLEDGMENTS

The authors thank the members of the Child Speech Research Lab for their help with data collection. A special thank you to the children and families for their participation. This work was sponsored by the Speech and Hearing Science Department and the Campus Research Board (#RB21001, PI: Mary Flaherty) at the University of Illinois Urbana-Champaign.

6. REFERENCES

[1] A. Astolfi, P. Bottalico, and G. Barbato, "Subjective and objective speech intelligibility investigations in primary school classrooms," *The Journal of the Acoustical Society of America*, vol. 131, no. 1, pp. 247–257, 2012.

[2] T. Houtgast, "The effect of ambient noise on speech intelligibility in classrooms," *Applied Acoustics*, vol. 14, no. 1, pp. 15–25, 1981.

[3] M. B. Kuhlenengel, L. C. Brill, S. Deng, H. F. Lester, J. Bovaird, J. Lau, L. M. Wang, and C. E. Waters, "An investigation of school environmental effects on student achievement," in *AEI 2017*, pp. 927–939, 2017.

[4] L. M. Wang and L. C. Brill, "Speech and noise levels measured in occupied k–12 classrooms," *The Journal of the Acoustical Society of America*, vol. 150, no. 2, pp. 864–877, 2021.

[5] N. Roy, R. M. Merrill, S. Thibeault, R. A. Parsa, S. D. Gray, and E. M. Smith, "Prevalence of voice disorders in teachers and the general population," *Journal of Speech, Language, and Hearing Research*, vol. 47, no. 2, pp. 281–293, 2004.

[6] L. J. Leibold and E. Buss, "Children's identification of consonants in a speech-shaped noise or a two-talker masker," *Journal of Speech, Language, and Hearing Research*, vol. 56, pp. 1144–1155, 2013.

[7] R. Y. Litovsky, "Speech intelligibility and spatial release from masking in young children," *The Journal of the Acoustical Society of America*, vol. 117, no. 5, pp. 3091–3099, 2005.

[8] J. E. Dockrell and B. M. Shield, "Acoustical barriers in classrooms: The impact of noise on performance in the classroom," *British Educational Research Journal*, vol. 32, no. 3, pp. 509–525, 2006.

[9] B. M. Shield and J. E. Dockrell, "The effects of environmental and classroom noise on the academic attainments of primary school children," *The Journal of the Acoustical Society of America*, vol. 123, no. 1, pp. 133–144, 2008.

[10] S. A. Stansfeld, B. Berglund, C. Clark, I. Lopez-Barrio, P. Fischer, E. Öhrström, M. M. Haines, J. Head, S. Hygge, I. Van Kamp, and B. F. Berry, "Aircraft and road traffic noise and children's cognition and health: A cross-national study," *The Lancet*, vol. 365, no. 9475, pp. 1942–1949, 2005.

[11] R. W. McCreery, M. Spratford, B. Kirby, and M. Brennan, "Individual differences in language and working memory affect children's speech recognition in noise," *International Journal of Audiology*, vol. 56, no. 5, pp. 306–315, 2017.

[12] R. W. McCreery, M. K. Miller, E. Buss, and L. J. Leibold, "Cognitive and linguistic contributions to masked speech recognition in children," *Journal of Speech, Language, and Hearing Research*, vol. 63, no. 10, pp. 3525–3538, 2020.

[13] M. Rudner, V. Lyberg-Åhlander, J. Brännström, J. Nirme, M. K. Pichora-Fuller, and B. Sahlén, "Listening comprehension and listening effort in the primary school classroom," *Frontiers in psychology*, vol. 9, p. 1193, 2018.

[14] R. L. Lewis, S. Vasishth, and J. A. Van Dyke, "Computational principles of working memory in sentence comprehension," *Trends in cognitive sciences*, vol. 10, no. 10, pp. 447–454, 2006.

[15] M. Klatte, T. Lachmann, and M. Meis, "Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting," *Noise and Health*, vol. 12, no. 49, p. 270, 2010.

[16] T. P. Hogan, S. M. Adlof, and C. N. Alonzo, "On the importance of listening comprehension," *International Journal of Speech-Language Pathology*, vol. 16, no. 3, pp. 199–207, 2014.

[17] J. R. Sullivan, H. Osman, and E. C. Schafer, "The effect of noise on the relationship between auditory

working memory and comprehension in school-age children," *Journal of Speech, Language, and Hearing Research*, vol. 58, no. 3, pp. 1043–1051, 2015.

[18] M. K. Pichora-Fuller, B. A. Schneider, and M. Daneman, "How young and old adults listen to and remember speech in noise," *The Journal of the Acoustical Society of America*, vol. 97, no. 1, pp. 593–608, 1995.

[19] American National Standards Institute, "American national standard specification for audiometers (ansi/asa s3.6-2018)." `https://webstore.ansi.org/Standards/ASA/ANSIASAS1260Part2010R2020`, 2018.

[20] M. Ross and J. Lerman, "A picture identification test for hearing-impaired children," *Journal of Speech and Hearing Research*, vol. 13, no. 1, pp. 44–53, 1970.

[21] D. V. M. Bishop, *Test for reception of grammar: Version 2: TROG-2*. London, UK: Pearson Assessment, 2003.

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