DOI: 10.61782/fa.2023.0062

EFFECTS OF SHORT-TERM IMPLICIT VOICE FAMILIARIZATION ON CHILDREN'S SPEECH-IN-NOISE RECOGNITION

Mary Flaherty[1*] **Rachael Price**[1]
[1] Department of Speech and Hearing Science, University of Illinois Urbana-Champaign, USA

ABSTRACT

The goal of this study was to evaluate the extent to which implicit voice familiarization via an interactive computer game can improve children's ability to recognize words presented in classroom noise. Children's (8-12 yrs.) closed-set word recognition was measured in a classroom noise masker before and after voice familiarization. Target words were spoken by two females, one the children would be familiarized with ("familiar talker"), and one they would not ("unfamiliar talker"). Following the pre-test, children were given a take-home computer game, designed for this study, to be played for 10 minutes a day for 5 days. During the game, children heard one voice ("familiar talker") instructing them to move their avatar towards various objects. Children then returned to the lab for the post-test. To evaluate the benefit of voice familiarity and test for generalization effects, four conditions were tested (pre- and post-gameplay): (1) familiar talker, familiar words, (2) unfamiliar talker, familiar words (3) familiar talker, unfamiliar words, and (4) unfamiliar talker, unfamiliar words. Working memory, attention, and inhibitory control were measured using the NIH toolbox cognition subtests. Preliminary results indicate word recognition improved for conditions with the familiar talker, highlighting the importance of knowledge-based factors during children's speech-in-noise recognition.

Keywords: *speech recognition, children, talker familiarity, implicit learning, familiarity training*

[*]***Corresponding author***: maryflah@illinios.edu

1. INTRODUCTION

It is well established that children are significantly poorer at understanding speech in the presence of background noise compared to adults [1-4]. This is despite a mature peripheral auditory system by the first year of life. This is believed to be due, in part, to children's immature sound segregation and selective attention, as well as their reliance on cue redundancy [5-6]. This suggests that children need years of listening experience and neural maturation to perform like adults when recognizing speech in noise. The focus of the current study is whether increasing listening experience with a specific voice via gameplay can improve children's speech recognition abilities. A child's age and listening experience are thought to play a critical role in the development of the skills needed to understand speech in noise. At the extreme end, absent or altered auditory input can interrupt or delay development of these processes, as in the case of early hearing loss [7-9]. However, individual experiences with specific auditory stimuli can also impact perception of that stimuli, including for listeners with typical hearing [10-13]. Experience with a talker's voice, like experience with a language, can also impact subsequent understanding of that voice. This is thought to be related to the listener's internal representations of the talker's voice characteristics, which become more consistent through exposure. For adults, this experience, or familiarity, with a talker's voice results in improved word and sentence intelligibility, improved recognition memory, and decreased processing time [14-15]. Repeated exposure to a voice, increases the fidelity of the internal representation of that talker's speech, leading the listener to better attend to and understand that talker's speech in degraded or noisy conditions, suggesting that familiarity can facilitate stream segregation and selective attention. Talker familiarity may be especially important for children, given their limited attention, slower processing speeds, and inability to rely on any single acoustic cue when making phonetic distinctions [16-20]. Considering that speech cues are further reduced by reverberation, this may be especially important for understanding speech compromised by poor classroom acoustics [21-22]. If children experience an advantage from hearing a familiar voice, the improvement in speech processing could free up cognitive resources for other tasks. Evidence shows that children can store information about a speaker to retrieve and use at a later time [23], similar to what has been found in adults [24-27], but research is limited regarding children's recognition of speech produced by familiar voices in noise.

The current study examines how experience with a talker's voice (talker familiarity) impacts children's speech recognition in typical classroom noise. Studies of familiar talkers have shown that listeners of all ages (infants to adults) understand speech in quiet better when it is spoken by a familiar talker than an unfamiliar talker [28-32]. Talker-specific information provides additional acoustic information (F0 and VTL, idiosyncratic articulation, prosodic/rhythmic patterns) that listeners might rely on to aid segregation and selective attention in the presence of competing sounds [33]. Explicit talker training in children has also been shown to improve word recognition [34]. However, implicit familiarity has received less attention despite this being how most listeners become familiar with a voice. Evidence from adults show they are able to identify voices at a more difficult SNR if they had been implicitly exposed to that talker. Moreover, adults were able to transfer this knowledge to novel words for the familiar talker [34], suggesting that implicit voice training can confer a familiarity benefit.

The current project used a one-group pretest-posttest experimental design to explore the impact of implicit voice exposure on children's ability to recognize words in classroom noise to determine if short-term implicit familiarity with a voice can improve their speech understanding. Implicit voice familiarization occurred via an interactive computer game. Before and after implicit familiarization, children completed a word recognition task, using a digitized version of the Word Intelligibility by Picture Identification (WIPI) test [35]. The target words were spoken by two different female talkers, one who children would be familiarized with in the game, and one who would remain unfamiliar. Implicit voice familiarization occurred by having children play a 10-minute interactive game once a day for five days. After five days, the same WIPI test was given to each child again to determine if hearing the voice during gameplay improved their word recognition. For comparison, the WIPI test conditions included (1) the familiar talker voice (heard during the game) and (2) an unfamiliar talker voice (not heard during the game). Based on recent findings in our laboratory that long term implicit familiarity with a talker (child's mother) leads to improvement in speech recognition [36] and evidence that short-term implicit voice familiarity impacts adults [29,34], we hypothesized that children would perform better for the words spoken by the target talker who was also heard during the game they played. We expected no differences between baseline and post-test for the unfamiliar talker, used as a control condition. To evaluate the effects of word familiarity and talker familiarity, children were tested on words they heard in the game and words not heard in the game. Best performance was expected in the condition with the familiar talker and familiar words.

2. METHODS

2.1 Participants

Twenty-four children (ages 8.1-12.9 years) were recruited from the local Urbana-Champaign area. Criteria for participation include (1) native speaker of English, (2) normal hearing bilaterally, with thresholds ≤20 dB HL for octave frequencies between 250 and 8000 Hz (ANSI, 2018), (3) no history of cognitive or language disorders, and (4) home access to a laptop or PC (for the game). All parents provided written informed consent and children provided verbal assent prior to participation. Children were paid $15/hour for their participation. The average study payment was $30 for successful completion of the study. If the child withdrew from or did not complete the study, they were paid for the portion completed.

2.2 Word Recognition Stimuli

Target stimuli were the four lists of words (25 monosyllabic words per list) from the WIPI test [35]. This is a closed-set test developed to assess speech discrimination of young children, adapted for the current study to be presented using a computer. Two adult female talkers recorded the target words. The female talkers were native English speakers with a standard American English dialect and were from the central Illinois area. Both talkers were 24 years old, and both had an F0 = 183 Hz (SD = 29). The speech was recorded at a sampling rate of 44,100 Hz in a sound booth using a high-quality condenser microphone (Shure-KSM42) mounted 6 inches from the talker's mouth. The targets were recorded with the carrier phrase "Go towards the" prior to each word to ensure equal stress on the words. After recording, all words were isolated and removed from the carrier phrase. Then, they were combined with one carrier phrase selected from the recordings. One carrier phrase was selected to ensure consistency among all the trials. Before the experiment, all stimuli were normalized to have the same long-term root-mean-square (RMS) amplitude and were listened to by the investigator and verified for clarity prior to use.

The masker was composed of classroom noise. Recordings were obtained online via open-source format (https://freesound.org/people/abcopen/sounds/166207/).

This noise was recorded in a primary school classroom and included children chatter and tables and chairs being moved around. The masker and target words were scaled to have equal amplitude. All words were then combined with the classroom noise masker using a custom script in Praat [38]to have a SNR of -10 dB. The SNR was chosen based on pilot testing.

2.3 Familiarization Game

Parents were given the game and game equipment to take home. This packet included over-the-ear headphones (Sony MDR-ZX110), a sound card (Behringer UCA222), and a USB drive with the game. USB-C adapters (Aukey USB) were also provided for participants using Apple computers. They were also given detailed instructions and a game log where participants tracked their progress. Participants used this equipment for all procedures involved in this study (word recognition and game). Listeners used their personal computer or laptop to play the game. When the game was opened on their computer, verbal instructions for the game were spoken by Talker 1. The target words used in the game are the same WIPI word lists (Lists 1 and 3) used in the word recognition test. These words were used with the exact carrier phrase as described above. Stimuli during the game were presented in quiet (no masker). The same pictures used in the word recognition test were used during the game with the caveat that only two non-rhyming pictures were presented together at any time (see Fig. 1). At the introduction and ending screen, engaging music was played but was never played at the same time as speech.

Figure 1. A screenshot taken from the familiarization game showing two images, two tokens, and the avatar that is controlled by the user.

2.4 Procedures

Table 1 provides an overview of study procedures, completed over a 7-day period. On Day 1, all participants came to the lab and completed the word recognition test, at which time all the voices and words were unfamiliar to them. The WIPI test was completed in the laboratory via Qualtrics (www.qualtrics.com). For this test, participants saw six pictures on the screen and were instructed to click on the word they heard. In this test, the child heard the carrier phrase "Show me the," followed by the target word spoken by a female talker. A feature of the WIPI test is that the pictures displayed on any trial are all phonetically similar by vowels. Therefore, while this was a closed set task, the response set required fine-grained acoustic detail to correctly identify the word compared to regular closed set tasks. The screen automatically advanced to the next trial after the participants responded. Performance was based on number of items correct. There were four conditions and each condition used one of the four WIPI lists. These conditions were chosen to evaluate the effects of the familiar talker and familiar words. All participants completed all four conditions. Condition order was counterbalanced across participants. Words within each condition were presented randomly. Prior to participation, there were two practice trials (with a voice not used in the main experiment) to ensure understanding of the task. All participants understood the task. Participants completed the word recognition test on Day 1, before playing the game, and then again on Day 7, after playing the game. Participants were instructed to wear the provided headphones and sound card while taking the test.

Table 1. Timeline of procedures for each participant.

Day 1	Day 2-6	Day 7
WIPI in lab (pretest)	Game play at home	WIPI in lab (posttest)
Talker 1 & 2 (Lists 1-4)	Talker 1 (List 1 & 3) 10 min./day	Talker 1 & 2 (Lists 1-4)
Closed set picture-pointing	Move avatar to collect tokens	Closed set picture-pointing
-10 dB SNR	Words in quiet	-10 dB SNR

After completing the baseline word recognition test, children were given the packet to take home containing the game, headphones, and a soundcard through which the game audio was routed. The child and their parent were given clear instructions for how to connect the equipment and play the game. The same headphones and soundcard were used in the word recognition tests. Participants were told they would be playing a short child-friendly computer game that was designed for the purposes of this study. Custom laboratory software was used to create the game. Talker 1 ("familiar talker") provided instructions at the onset of the game and indicated actions to take during the game on each level with the intention to familiarize the listener with that voice. Participants were instructed to play two levels of the game once per day for five days. The same pictures used in the word recognition test were used during the game with the caveat that only two non-rhyming pictures were presented together at any time (see Fig 1). During the game, the participant used computer keys to move towards images corresponding to those words spoken by Talker 1 (e.g., "Go towards the chair"). Each time the participant moved toward the correct picture, a token was given to the participant, and the game continued. If the participant did not go toward the correct picture, no token was given, and the game continued. The goal of the game was to collect as many tokens as possible. The game continued for approximately 10 minutes (5 minutes per level). At the end of each level, the participant was given their token total. The game ended after all the words had been spoken, regardless of the participant's performance. Once started, the game was locked until the participant completed both levels, and thus participants could not end the game prematurely. Participants (or their parents) kept a log of the gameplay using the "Jungle Jogger Game Log". The purpose of this game was for the child to hear/be exposed to the person's voice guiding the game. They completed the word recognition test the day after the fifth training day (Day 7). Five days of training were chosen based on previous data showing that this amount of training is sufficient to elicit the talker familiarity advantage in quiet [39].

Given that working memory and inhibitory control are among the key cognitive abilities thought to support speech processing in noise [40-41], and to account for individual differences in performance, all children completed two tablet-based NIH Toolbox Cognition Battery subtests after the final word recognition test. These were (1) List Sorting Working Memory test (LSWM), and (2) Flanker test of Selective Attention and Inhibitory Control. Standard scores were used to represent each construct in the statistical models to evaluate their impact on pre- and post-gameplay word recognition scores.

3. RESULTS

A total of 24 participants were recruited for the study. However, one participant was mistakenly tested on the incorrect pre-test due to experimenter error, and another participant was unable to return for post testing due to his mother's illness. Therefore, they did not complete the study and their data were excluded from the analyses. The data from 22 participants was analyzed.

When first examining baseline performance, recognition scores across all four lists appeared to be roughly equal. The percent correct baseline scores (out of 25) for Condition 1: 73% (SD=8), Condition 2, 71.1% (SD=7), Condition 3, 70% (SD=11), and Condition 4, 71% (SD=9). When considering performance following gameplay, the largest improvement appeared to be in Condition 1, which had both the talker and the words heard during game play (Familiar Talker/Familiar Words; see Fig. 2). The mean improvement for this condition an increase of 11%. The mean improvement in Condition 2 (Unfamiliar Talker/Familiar Words) was 6%, Condition 3 (Familiar Talker/Unfamiliar Words was 1%, and Condition 4 (Unfamiliar Talker/Unfamiliar Words) was 4%. In other words, out of 25 possible points, the mean improvement for Condition 1 was 2.8 points (SD 2.8), Condition 2 was 1.5 points (SD 2.4), Condition 3 was 0.4 point (SD 3.5), and Condition 4 (Unfamiliar Talker/Unfamiliar Words) was 1.2 points (SD 2).

To examine the statistical significance of these trends, generalized linear models (GLMM) were fit using the listeners' binary responses (correct or incorrect recognition score on each trial). The independent variables included the List Condition, timepoint (before or after gameplay), and child age. Two models (pre- and post-gameplay) were used to analyze the effects of talker familiarity and word familiarity on speech recognition, including condition (List 1-4) and (age).

When considering the baseline, pre-gameplay conditions, the model showed no statistically significant difference between the lists, $p > 0.10$. The model did find a statistically significant difference in the post training conditions. Children were able to correctly recognize 84% of words of list 1 (familiar talker, familiar words), 78% of words of list 2 (unfamiliar talker, familiar words), 70% of words of list 3 (familiar talker, unfamiliar words), and 75% of list 4 (unfamiliar talker, unfamiliar words). Tukey's post hoc comparisons for the post-gameplay conditions showed a statistically significant difference in speech recognition between lists 1 and 2 (X^2=-0.40, p=0.047), lists 1 and 3 (X^2=-0.80, p<0.001), lists 1 and 4 (X^2=-0.52, p=0.003), and lists 3 and 2 (X^2= -0.40, p=0.021), wherein list 1 scores were better than lists 2-4, and list 3 scores were also lower than list 2.

Child age also showed a significant effect on the pre-gameplay (baseline) conditions, such that performance was better for older children compared to younger children (X^2= -0.008, p=0.007). However, the model for the post training (figure 2B) condition did not show any effect of age, with the exception of List 4 (unfamiliar talker, unfamiliar words), wherein performance increased as the age increased (X^2= -0.06, p=0.008).

Two models were also used to analyze the effect of WM and IC on speech recognition scores for the pre and post gameplay conditions. Both models showed no statistically significant effects of WM and IC on speech recognition.

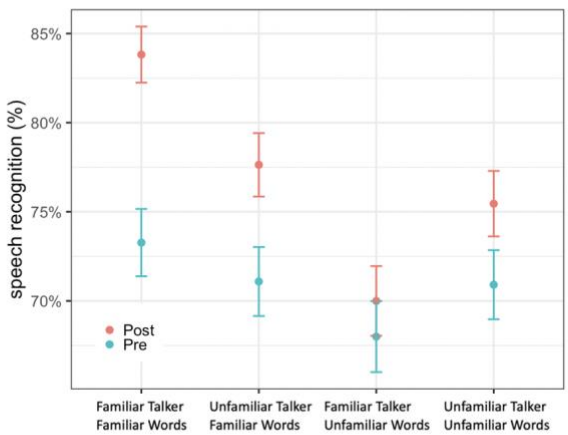

Figure 2. Pre- and post-gameplay speech recognition scores (% correct) for each condition.

4. DISCUSSION.

The goal of the current study was to determine if implicit familiarization with a voice via an interactive game can improve word recognition in noise for children. Prior to playing the game, children showed similar baseline word recognition performance across all test conditions, indicating similar difficulty across conditions. After playing the game for 5 days, totaling 50 minutes of voice exposure, mean scores improved only in the test condition with the familiar talker and the familiar words. These results suggest that while short-term implicit familiarization may lead to improvement in children's word recognition, the advantage is does not generalize to unfamiliarized words. For the conditions with the unfamiliar voice, there no significant improvement after playing the game when only the target words were familiar (Condition 2) or when the target words were unfamiliar (Condition 4). The lack of improvement for these conditions suggests that children relied on familiarity with both the voice and the words, and not just the words alone. There was also no significant improvement in word recognition performance when the words and voice were both unfamiliar, suggesting that the improvement observed in Condition 1 was not due to practice effects from taking the test at two time points, as this would be expected to be evident across all conditions. However, the absence of improvement for the condition with the familiar words and the unfamiliar voice indicates that children only benefit from the combination of having experience with both the words and the talker, but not one or the other. This may suggest that children's internal representations of the talker's voice characteristics do become more consistent through exposure, but that effects of this exposure are only evident when those representations match directly to what they have been exposed to. This could suggest that both talker and word familiarity combined may facilitate word recognition, possibly by improving stream segregation and selective attention for the familiarized stimuli.

The lack of generalization to words not played in the game was not consistent with our predictions. Prior studies in adults have shown that familiarity with a voice improves perception and this improvement generalizes to novel words. [42-43]. This does not appear to be true for children and brings into question whether there is a developmental effect in generalizability from talker familiarity. While there is very limited research on talker familiarity effects in children, at least one prior study showed similar effects as our current findings in a group of infants. Infants listened significantly longer to words in noise spoken by their own mother, as long as the infants had been familiarized with the words [28]. That is, familiarity effects for their mother's voice were not observed for novel words. The current study extends those findings to school-age children with a voice that they have had much less exposure with compared to their experience with their own mother's voice.

Studies of toddlers have shown that they are able to generalize talker-specific knowledge to novel words following brief exposure to accented speech, but this was when perceiving speech sounds in quiet. The current work suggests that adding noise might impact children's ability to generalize their learning. The added difficulty resulting from the noise may force the child to rely on familiarity with multiple sources in order to obtain an advantage. This would be consistent with research suggesting children rely more heavily on access to redundant acoustic speech cues when recognizing speech in the presence of other sounds. Speech recognition in noise is more difficult for children than for adults, and this is thought to be due not only to immature segregation and selective attention but also an inability to recognize speech based on sparse cues [5]. It may be that in noise, children in the current study were able to utilize the internal representations only if they were fully consistent with the speech heard in game. Additional considerations of word familiarity and generalization effects will be examined in future studies. Given that all words in the study were familiar to the children, baseline word familiarity is not believed to have played a role in the findings. However, additional exploration into the effects of noise and listener age on these processes is needed. Moreover, additional exposure to the voice may be needed to observe a generalizable effect on speech recognition. The current manipulation of only 50-minutes of exposure may have been too little to have shown an effect, especially in our small sample. Future directions will explore these findings using sentence stimuli and open-set tasks to avoid confounding factors associated with becoming familiar with the pictures used in the task.

The findings related to child age and cognition were counter to our predictions. The effects of child age were expected to impact performance in both pre- and post-test conditions, based on prior data [1-6]. However, age effects were only present at baseline, and were no longer observed following gameplay, except for the condition with both an unfamiliar talker and unfamiliar words. This suggests that short-term exposure to either a talker or a list of words may lead to a reduction in age effects. This could indicate a potential interaction between age, recognition, and familiarity, but additional work is needed to tease apart this finding. In addition, although the sample size was relatively small, the results indicated that neither working memory nor selective attention was predictive of performance. This was observed across all timepoints, suggesting that these two tests do not assess the processes responsible for individual differences in speech recognition in the present study, nor the ability to benefit from short-term implicit talker familiarity.

In conclusion, the findings from this preliminary investigation indicate that short-term implicit talker familiarity can impact children's speech recognition, highlighting the importance of knowledge-based factors during children's speech-in-noise recognition. Children's word recognition with a specific voice improved after being exposed to that voice via an interactive computer game played for 10-minutes a day for 5-days, at least for the words they heard that voice speak in the game. The lack of generalization to words not heard in the game could indicate children may rely more on a combination of acoustic familiarity cues when recognizing speech in noise.

5. ACKNOWLEDGMENTS

The authors thank the members of the Child Speech Research Lab that helped with data collection. A special thank you to Colton Price for contributing to computer game development and design. This work was sponsored by the Speech and Hearing Science Department of the University of Illinois at Urbana-Champaign.

6. REFERENCES

[1] Corbin, N.E., Bonino, A.Y., Buss, E., Leibold, L.J. (2016). Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. Ear Hear. 37, 55-63. https://doi.org/10.1097/AUD.0000000000000201

[2] Flaherty, M.M., Buss, E., Leibold, L.J. (2019). Developmental effects in children's ability to benefit from F0 differences between target and masker speech. Ear Hear. 40, 927-937. https://doi.org/ 10.1097/AUD.0000000000000673

[3] Hall, J.W., Grose, J.H., Buss, E., Dev, M.B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children.

Ear Hear. 23, 159-165. https://doi.org/ 10.1097/00003446-200204000-00008

[4] Leibold, L.J., Buss, E. (2013). Children's identification of consonants in a speech-shaped noise or a two-talker masker. J. Speech Lang. Hear. 56, 1144-1155. https://doi.org/10.1044/1092-4388(2012/12-0011

[5] Leibold, L. J., & Buss, E. (2019). Masked speech recognition in school-age children. *Frontiers in Psychology*, *10*(September). https://doi.org/10.3389/fpsyg.2019.01981

[6] Flaherty, M. M., Buss, E., & Leibold, L. J. (2021). Independent and combined effects of fundamental frequency and vocal tract length differences for school-age children's sentence recognition in a two-talker masker. *Journal of Speech, Language, and Hearing Research*, *64*(1), 206–217. https://doi.org/10.1044/2020_JSLHR-20-00327

[7] Sharma, A., Dorman, M. F., Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. *Ear and Hearing*, *23*, 532–539

[8] Yoshinaga-Itano, C., Sedey, A. L., Coutler, D. K., et al. (1998). Language of early-and later-identified children with hearing loss. *Pediatrics, 5*, 1161–1171.

[9] Moeller, M. P., Tomblin, J. B., & the OCHL Collaboration. (2015). Conclusions and implications for research and practice. *Ear and Hearing*, *36*, 92–98. https://doi.org/10.1007/978-3-8349-9797-5

[10] Moore, D. R. (2002). Auditory development and the role of experience. *British Medical Bulletin*, *63*(1), 171–181. https://doi.org/10.1093/bmb/63.1.171

[11] Skoe E, Kraus N. Auditory Brain Stem Response to Complex Sounds: A Tutorial. *Ear and Hearing*, 31, 302–324. [PubMed: 20084007]

[12] Boothroyd A. (2010). Adapting to changed hearing: the potential role of formal training. *Journal of the American Academy of Audiology*, *21*(9), 601–611. https://doi.org/10.3766/jaaa.21.9.6

[13] Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics, 27*(4), 359–384

[14] Ladefoged, P., & Broadbent, D. E. (1957). Information Conveyed by Vowels. *The Journal of the Acoustical Society of America*, *29*(1), 98–104. https://doi.org/10.1121/1.1908694

[15] Laver, J. (1989). Cognitive science and speech: A framework for research. In H. Schnelle & N. O. Bernsen (Eds.),*Logic and linguistics* (Research Directions in Cognitive Science: European Perspectives, Vol. 2, pp. 37–70). Hillsdale, NJ: Erlbaum.

[16] Laver, I., & Trudgill, P. (1979). Phonetic and linguistic markers in speech. In K.R. Scherer & H. Giles (Eds.), *Social Markers in Speech* (pp. 1-32). Cambridge, England: Cambridge University Press

[17] Van Lancker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition: Patterns and parameters, Part I: Recognition of backward voices. *Journal of Phonetics*, 13. 19-38

[18] Clarke, C.M., Garrett, M.F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America, 116* (6), 3647e3658. http://doi.org/10.1121/1.1815131.Theodore et al., 2015

[19] Idemaru K, Holt LL. (2013) The developmental trajectory of children's perception and production of English/r/-/l. *The Journal of the Acoustical Society of America, 133*(6):4232–46.

[20] Mayo, C., & Turk, A. E. (2003). Is the development of cue weighting strategies in children's speech perception context-dependent? *International Congress of Phonetic Sciences*, 677–681.

[21] Hazan, V., Barrett, S. (2000) The development of phonemic categorization in children aged 6–12. *Journal of Phonetics*, *28*(4), 377–96.

[22] Bronkhorst, A.W. (2015) The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Attention Perception and Psychophysics*. 77(5):1465–87.

[23] Peng, E., Easwar, V. (2021) Development of temporal cues perception in noise and reverberation. *Journal of the Acoustical Society of America,* 150(4): A337–8.

[24] Astolfi, A, Bottalico P, Barbato G. (2012) Subjective and objective speech intelligibility investigations in primary school classrooms. *The Journal of the Acoustical Society of America, 131*(1):247–57.

[25] Bradley, J.S. Speech intelligibility studies in classrooms. *The Journal of the Acoustical Society of America,* 80(3):846–54.

[26] Levi, S. V. (2015). Talker familiarity and spoken word recognition in school-age children. *Journal of Child Language*, *42*(4), 843–872. https://doi.org/10.3109/10409238.2016.1143913.PP2A

[27] Levi, S. V., Winters, S. J., & Pisoni, D. B. (2011). Effects of cross-language voice training on speech perception: Whose familiar voices are more intelligible? *The Journal of the Acoustical Society of America*, *130*(6), 4053–4062. https://doi.org/10.1121/1.3651816

[28] Barker, R. G., & Newman, J. P. (2004). Handbook of personology and psychopathology. Wiley.

[29] Newman, J. P., & Evers, K. E. (2007). Theories of the self and psychopathy: A reconceptualization. Clinical Psychology Review, 27(1), 58-72. https://doi.org/10.1016/j.cpr.2006.09.005

[30] Magnuson, J. S., Mirman, D., & Luthra, S. (2021). Auditory word recognition: Evidence from the auditory lexical decision task. Journal of Memory and Language, 116, 104217. https://doi.org/10.1016/j.jml.2020.104217

[31] Souza, P. E., Arehart, K. H., Kates, J. M., Croghan, N. B. H., & Gehani, N. (2013). Exploring the relationship between working memory and hearing aid benefit. Hearing Research, 304, 26-33. https://doi.org/10.1016/j.heares.2013.06.005

[32] Tye-Murray, N., Spehar, B., Myerson, J., Hale, S., & Sommers, M. S. (2016). Reading benefits and costs for children with and without dyslexia. Journal of Experimental Child Psychology, 146, 127-146. https://doi.org/10.1016/j.jecp.2016.01.002

[33] Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. Perception & Psychophysics, 60(3), 355-376. https://doi.org/10.3758/BF03206093

[34] Kreitewolf, J., Friederici, A. D., & von Kriegstein, K. (2017). Hemispheric lateralization of linguistic prosody recognition in comparison to speech and speaker recognition. NeuroImage, 152, 218-229. https://doi.org/10.1016/j.neuroimage.2017.02.001

[35] Ross, M., & Lerman, J. (1970). An experimental investigation of the resolution of ambiguity by young children. Child Development, 41(2), 407-422. https://doi.org/10.2307/1127385

[36] Flaherty, M. (2020) The effect of talker familiarity on children's speech-in-speech recognition. Poster presented at American Auditory Society Scientific and Technology Meeting, Scottsdale, AZ, March 2020.

[37] ANSI, 2018. American National Standard Specifications for Audiometers. ANSI S3.6-2018. New York.

[38] Boersma, P., & Weenink, D. (2021). Praat: Doing phonetics by computer [Computer program]. Version 6.1.43. https://www.fon.hum.uva.nl/praat/

[39] Levi, S. V., Ibrahim, R. K., & Fishman, G. Y. (2019). Perception of non-native tonal contrasts: A multidimensional scaling analysis. Journal of Phonetics, 77, 100923. https://doi.org/10.1016/j.wocn.2019.100923

[40] McCreery, R. W., Miller, M. K., Buss, E., & Leibold, L. J. (2020). Cognitive and linguistic contributions to masked speech recognition in children. *JSLHR* *63(*10), 3525–3538. https://doi.org/10.1044/2020_JSLHR-20-00030.

[41] Foo, C., Rudner, M., Rönnberg, J., & Lunner, T. (2007). Recognition of speech in noise with new hearing instrument compression release setting requires explicit cognitive storage and processing capacity. *Journal of the American Academy of Audiology, 18*, 618-631.

[42] Allen, J. S., & Miller, J. L. (2004). Factors affecting perceptual learning of speech: Evidence from the processing of a phonetic contrast. *Journal of the Acoustical Society of America,* 115(3), 1318-1330. https://doi.org/10.1121/1.1635849

[43] Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. Perception & Psychophysics, 67(2), 224-238. https://doi.org/10.3758/BF03195277