Word perception errors of children and adults in noise and quiet

Anna Sfakianaki
University of Ioannina, Greece

https://doi.org/10.36505/TheLinguisticProceedings/2025/17/02/018/000704

Abstract

Accurate speech perception is vital for spoken language comprehension and effective communication. Especially for children, spoken word recognition is crucial for vocabulary, language, and knowledge acquisition in general. However, listening conditions in the classroom are often adverse, due to poor classroom acoustics. The present study investigates word perception errors made by adults and children in noise and quiet. Besides performance (word recognition score), word misconceptions are also analysed in relation to phonological, semantic and other lexical properties, i.e. word familiarity and frequency as well as phonological neighbourhood density metrics. This type of error analysis can be more informative of the nature of difficulties encountered by the listener during speech recognition in adverse conditions, and provide insights into the relationship between linguistic knowledge and speech perception in noise.

Keywords: speech perception, noise, word errors, children, adults

Introduction

Previous research has shown that children are more susceptible to auditory masking than adults (Leibold, Buss 2019; Sfakianaki et al., 2021). Although quantitative measures, such as percent correct word or phoneme scores, have been primarily employed to explore differences in child vs adult performance, a qualitative analysis of the types of errors children make in comparison with those made by adults would help more accurately define the difficulties faced when identifying speech in noise. The current study investigates errors made by adults and children with normal hearing and typical development in the perception of Greek disyllabic words, without the aid of context, in quiet and in two-talker noise, which has been found more challenging than speech-shaped noise (Corbin et al. 2016). The analysis focuses on the effect of noise on word misperceptions in relation to phonological, semantic and other lexical properties. Such properties have been examined in prior studies, and have been found to play a role -albeit not always straightforward, in speech perception speed and accuracy (e.g. Braza et al. 2022; Garlock et al. 2001).

Methodology

A corpus was created for the experiment, consisting of disyllable words embedded in the phrase "Say the word …". The corpus included 250 familiar

words selected from the Greek first-grade primary school textbook and 250 low frequency words, expected to be unfamiliar to the children, selected from database Greeklex 2 (Kyparissiadis et al. 2017). The corpus was recorded with one female adult speaker with no speech or hearing disorders. The clean stimuli were then mixed with a two-talker masker, created by superimposing the separate recordings of two female adult speakers reading children's literature excerpts. Subsequent to a pilot experiment, the SNR level for adults was set at 2.5 dB and for children at 7.5 dB, so as to avoid floor or ceiling effects.

Ten adults (mean age: 36.3, SD: 12.35) and ten children (mean age: 7.08, SD: 0.31), who had just completed the first year of primary school, participated in the perception experiment. Participants were Greek native speakers of typical development with no speech or hearing disorders, subsequent to thorough audiometric evaluation. Each participant listened to 80 randomized stimuli, 40 in noise and 40 in quiet, at 65 dB SPL from a loudspeaker placed in front of the listener at a radial distance of 0.9m in a soundproof recording studio.

Responses that deviated in any way from the original lexical form were classified as incorrect. The binary score data amounted to a total of 1600 items and was analyzed using Mixed Analysis of Variance (fixed factors: age group, gender, noise, word familiarity; random factors: participant, word). An error analysis of the word misconceptions in noise followed. A comparison between target and recognized word was conducted regarding 6 types of measures: 1) phonemic alteration type (substitution, omission, insertion or a combination of alterations), 2) phonemic alteration position in word (first syllable, second syllable, both syllables or addition of syllable(s)), 3) semantic alteration, 4) word frequency alteration, 5) Levenshtein distance, a comparison of target word sequence with recognized word sequence, calculating a single distance, and 6) PLD20 (Phonological Levenshtein Distance), a summary statistic over the word's 20 nearest neighbours in order to show phonological neighbourhood density difference between target and recognized word.

Results

Regarding the binary score analysis, both children and adults showed statistically worse performance in noise vs quiet and in low vs highly familiar words in noise, but not in quiet. Age (df= 1, F=0.05, p=0.834) and gender (df=1, F=0.18, p=0.676) were not significant factors, but noise (df=1, F=660.36, p=0.000) and word familiarity (df=1, F=59.95, p=0.000) were statistically significant, along with the random factor 'word' (p=0.000). Interactions age*noise (df=1, F=5.26, p=0.022) and noise*familiarity (df=1, F=70.19, p=0.000) were also found statistically significant.

A total number of 256 word misconceptions in noise were made by both age groups. Children and adults produced the same number of misconceptions when SNR level was 5 dB higher for children in comparison with that for adults. Also,

errors in unfamiliar words were twice as many as in familiar words in both age groups. Regarding the word error analysis, the following observations were made:

1) *Phonemic alteration type:* Omission (adults and children: 5.47%) and insertion (adults: 11.72%, children: 18.75%) occurred less frequently for both age groups, while substitution (adults: 28.13%, children: 39.06%) or a combination of the above alterations (adults: 54.69%, 35.94%) were observed more frequently.

2) *Phonemic alteration position*: Phonemic alteration occurred more frequently across both syllables for adults (34.38%, children: 20.31%), while children more often altered the first syllable (37.50%, adults: 22.66%) or added syllables either before or after the target word, especially in familiar words (33.33%, adults: 23.40%).

3) *Semantic alteration*: For both age groups the recognized word was more often semantically different than the target word (adults: 81.25%, children: 60.16%); child misconceptions were more often nonexistent words, especially when the target was an unfamiliar word (40.70%, adults: 13.58%).

4) *Word frequency alteration*: For both age groups, the recognized word often had higher word frequency than the target word, which was more evident in words of lower familiarity (adults: 70.37%, children: 50.00%).

5) *Levenshtein distance*: According to independent-samples t-test, the Levenshtein distance between target and recognized word was significantly higher for adults (M: 2.47, SD: 1.48) than children (M: 2.05, SD: 1.20) ($t(256)=2.51$, $p=0.013$). It was also higher in familiar (M: 2.42, SD: 1.19) vs unfamiliar (M: 2.17, SD: 1.44) words, but the effect was not statistically significant ($t(256)=1.52$, $p=0.13$). 6) *Phonological Levenshtein Distance*: Based on paired-sample t-tests, PLD20 of the recognized word (M: 1.70, SD: 0.47) was not significantly higher from PLD20 of the target word for adults (M: 1.67, SD: 0.28, $t(128)=0.62$, $p=0.538$). In contrast, a statistically significant increase in PLD20 between the target (M: 1.63, SD: 0.29) and the recognized word was observed for children (M: 1.79, SD: 0.55, $t(128)=2.92$, $p=0.004$).

Discussion

Based on the above analyses, noise poses a greater obstacle in word recognition for children, in line with the literature (Leibold & Buss, 2019), as first graders required a 5 dB SNR advantage in two-talker noise in order to reach adult performance. Word recognition was lower for unfamiliar words in noise, but not in quiet, both for children and adults. Also, most lexical properties of misconceptions were significantly affected by word familiarity; for example, the word frequency increase and Levenshtein distance decrease in misconceptions were more pronounced for unfamiliar vs familiar word targets. Misconception analysis results indicate that, when struggling to recognize a word in noise, children do not hesitate to venture words that do not exist and that their misconceptions are closer phonologically (lower Levenshtein distance) to the

target word as compared with adults. Similar findings have been reported for the misconceptions of younger vs older adults (Vickery, 2021). Children seem to be more constrained by the actual sounds of the target word; they report a higher percentage of phonemes as heard despite the fact that, in the end, the recognized word does not make sense or is less frequent. Adults seem to pay more attention to meaning and to the frequency of occurrence of a word due to their linguistic experience and richer vocabularies. In addition, when striving to recognize words in noise, especially unfamiliar ones, both children and adults report words higher in frequency of occurrence, as also documented in previous work (e.g. Cooke et al., 2019 for adults), but children also seem to select words with far less phonological neighbours (high PLD20). Thus, under perceptual stress, children tend to compensate for weak lexical knowledge by relying on phonological distinctiveness.

Acknowledgements

Special thanks are due to Dr. George Kafentzis for assistance with the experimental setup, and Dr. Devora Kiagiadaki for the audiological evaluation of participants.

References

Braza, M. D., Porter, H. L., Buss, E., Calandruccio, L., McCreery, R. W., Leibold, L. J. 2022. Effects of word familiarity and receptive vocabulary size on speech-in-noise recognition among young adults with normal hearing. PloS one, 17(3), e0264581. https://doi.org/10.1371/journal.pone.0264581

Cooke, M., García Lecumberri, M. L., Barker, J., Marxer, R. 2019. Lexical frequency effects in English and Spanish word misperceptions. JASA, 145(2), EL136.

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 and Hearing, 37(1), 55-63. https://doi.org/10.1097/AUD.0000000000000201

Garlock, V. M., Walley, A. C., Metsala, J. L. 2001. Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. J. Mem. Lang., 45(3), 468–492.

Kyparissiadis, A., van Heuven, W.J.B., Pitchford, N.J., Ledgeway, T. 2017. GreekLex 2: A comprehensive lexical database with part-of-speech, syllabic, phonological, and stress information. PLoS ONE 12(2), e0172493.

Leibold, L. J., Buss, E. 2019. Masked speech recognition in school-age children. Frontiers in Psychology, 10, 1981. https://doi.org/10.3389/fpsyg.2019.01981

Sfakianaki, A., Kafentzis, G., Kiagiadaki, D., Vlahavas, G. 2021. Effect of face mask and noise on word recognition by children and adults. Proceedings ExLing 2021, 217-220. https://doi.org/10.36505/ExLing-2021/12/0055/000528

Vickery, H. B. 2021. Lexical properties of perceptual errors made by younger and older adults listening to speech in multitalker babble. (Master's thesis). Retrieved from https://scholarcommons.sc.edu/etd/6486