DO NONNATIVE LISTENERS BENEFIT AS MUCH AS NATIVE LISTENERS FROM SPATIAL CUES THAT RELEASE SPEECH FROM MASKING?

Avivi-Reich, M. and Schneider, B.A.
Department of Psychology, University of Toronto Mississauga, Canada
Centre for Research on Biological Communication Systems, University of Toronto Mississauga
E-mail: meital.avivi@utoronto.ca

Abstract

Nonnative listeners generally perform worse than native listeners on measures of speech perception, especially in acoustically complex environments. An important question that arises is whether nonnative listeners can benefit from acoustic cues that aid speech perception to the same extent as native listeners. In recent studies we have examined the ability of native and nonnative listeners to take advantage of one such acoustic cue, namely spatial separation, when listening to a talker in a noisy environment. In these studies participants were asked to repeat semantically anomalous English sentences masked by either two-talker speech or speech-spectrum noise. The results indicate that even though non-natives require higher signal-to-noise ratios (SNRs) to recognize speech, they reap the same advantage from spatial separation as native speakers.

Non-native listeners generally perform worse than native listeners on measures of speech perception, especially in acoustically complex environments. An important question that arises is whether nonnative listeners can benefit as much as native listeners from the bottom-up (e.g., spectral separation of competing sounds, asynchronous fluctuations in the sounds' amplitude) and top-down (e.g., familiarity with a talker's voice or a priori knowledge of target location) cues that might aid speech perception in noisy environments. It is possible that in acoustically cluttered environments where bottom-up cues are likely to be obscured or masked by other sound sources, the reduced ability of nonnative listeners to extract phonetic information from target sentences might prevent them from taking full advantage of cues that could have aided them in their efforts to perceive the target speech. Hence, reduced ability to segregate the target speech from the background noise, could lead to sub-optimal linguistic and semantic processing when the incoming target is in one's second language.

In a previous study (Ezzatian, Avivi and Schneider, 2010) we examined the ability of native and nonnative listeners to take advantage of one such bottom-up acoustic cue, namely perceived spatial separation, when asked to listen to and repeat semantically anomalous English sentences masked by either two-talker speech or speech-spectrum noise.

Given previous findings regarding the importance of age of language acquisition and language competence in how well nonnative listeners perform on speech perception tasks, we recruited nonnative listeners who varied in their experience with the English language (from very little to extensive experience), and administered a vocabulary and reading comprehension task to all participants in order to assess language competence.

A total of 64 younger adults (18–26 years old) with normal hearing participated in this study. These individuals were subdivided into 4 groups according to their exposure to and experience with the English language: native speakers of English (n = 16), those who had learned English between 6 and14 years of age (n = 16), those who had learned English after 15 years of age (n = 16), and a mixed group (n = 16), consisting of individuals who had learned English before the age of 5, but in a country where English was not the native language (e.g., India or the United Arab Emirates).

Targets were semantically anomalous sentences with 3 keywords (e.g., "A house should dash to the bowl," (Helfer, 1997) spoken by a female talker. These sentences were presented against either two-talker anomalous speech (two females) or speech-spectrum noise. Half of the sentences were played while the participants perceived the target and masking sounds as emanating from the same spatial location (45^0 to the left of the listener) and half while the participants perceived the target and masking sounds as emanating from different spatial locations (target 45^0 to the left of the listener, masker 45^0 to the right of the listener) using the precedence effect. Targets and maskers were presented at 4 different signal-to-noise ratios; and for each individual, the dB SNR corresponding to his or her 50%-correct performance threshold was computed from their individual psychometric functions. Figure 1 shows the average performance of the participants from the four language competence groups in two types of maskers (speech spectrum noise vs two-talker speech), with (solid symbols) and without (open symbols) perceived spatial separation.

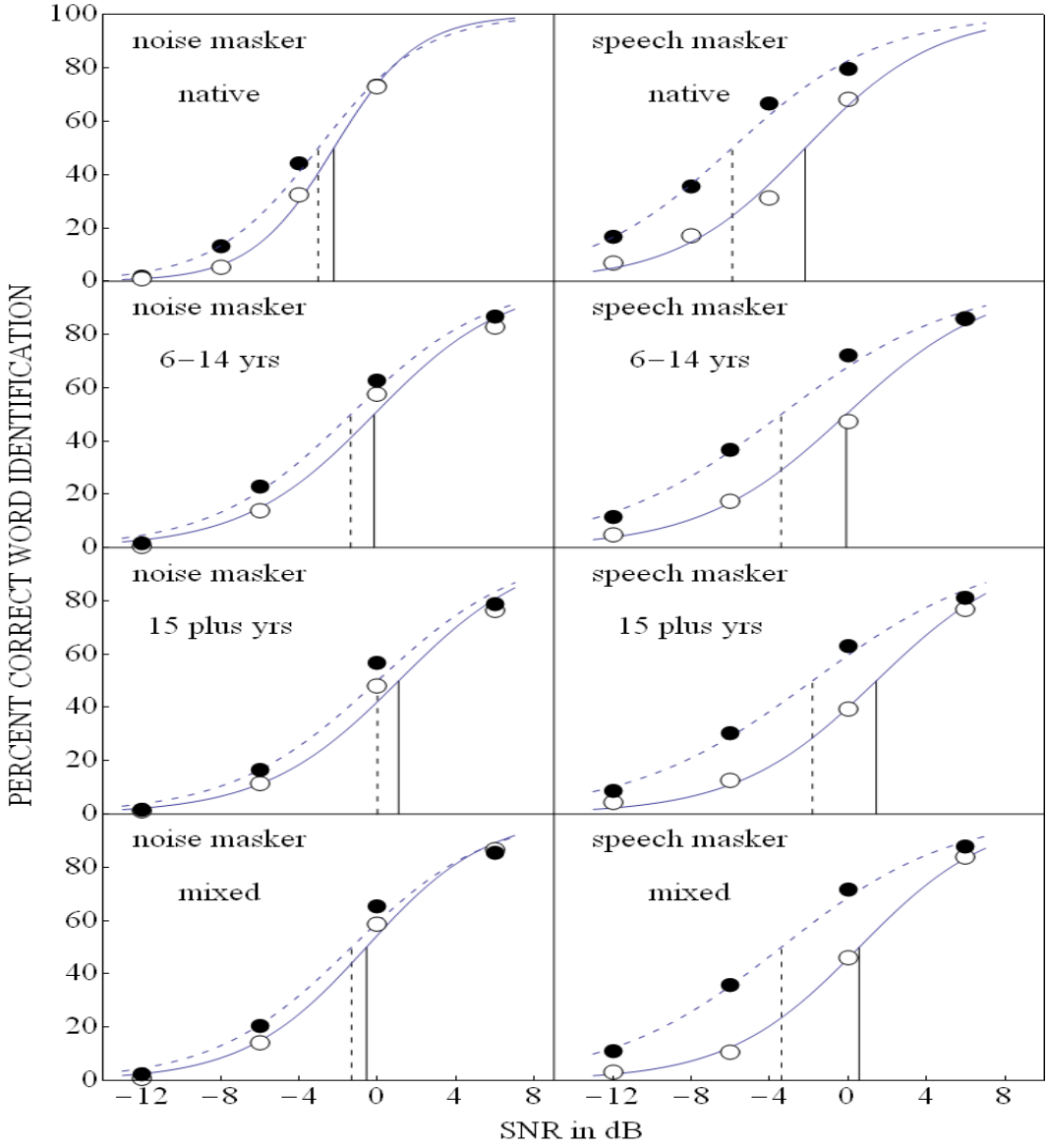

Fig. 1. Average percent-correct performance at different levels of SNR as a function of Group and Masker type, with separate lines for the co-located and perceptually separated conditions. From Ezzatian et al. (2010).

Figure 1 shows that native listeners, over all conditions, had lower thresholds (by 2-3 dB) than the nonnative listeners. Among the nonnative listeners, those in the 6-14 and mixed

groups had equivalent thresholds, which were lower than those in the 15+ years group by 1.4-1.5 dB. Figure 1 also shows that there was an advantage of spatial separation that was larger when the masker was speech (3.57 dB) than when the masker was noise (1 dB), but that nonnative listeners were not disproportionately affected by either type of masker, and derived the same benefit from spatial separation as native listeners under all conditions. Hence with respect to thresholds, the acoustic factors that release speech from masking are not affected by linguistic competence. However, Figure 1 shows that when speech was masked by noise the slopes of the psychometric functions were shallower for nonnatives than natives. But no such differences in slopes were found when speech was masked by speech. Hence linguistic competence does affect the slope of the psychometric function, but only when the masker is noise.

This study reflected the performance of nonnative listeners who immigrated to Canada from 22 countries: Brazil (n = 1), China (n = 8), Egypt (n = 1), Germany (n = 4), India (n = 6), Indonesia (n = 2), Israel (n = 1), Japan (n = 1), Korea (n = 2), Kuwait (n = 1), Lithuania (n = 1), Malaysia (n = 1), The Netherlands (n = 2), Nigeria (n = 1), Pakistan (n = 5), Russia (n = 1), Saudi Arabia (n = 1), Singapore (n = 2), Sri-Lanka (n = 2), Taiwan (n = 1), Ukraine (n = 1), and the United Arab Emirates (n = 2). When collapsing the results across all different languages, with the exception of slope differences in noise, native and nonnative listeners benefit equally from perceived spatial separation. However, there is evidence that a person's native language can affect the degree to which there is release from spatial masking.

A study (Wu et al., accepted) which was conducted both at the University of Toronto and at Peking University compared the degree to which monolingual speakers of English, listening to semantically anomalous English sentences, and monolingual speakers of Mandarin Chinese, listening to semantically anomalous Chinese sentences, benefitted from perceived spatial separation between the location of the target voice and the location of one of three different kinds of competing sources (a speech spectrum noise; anomalous sentences in the same language spoken by two other talkers; and anomalous sentences in the other language spoken by two other talkers).

Interestingly, although both English and Chinese listeners benefitted to the same extent from spatial separation when the masking sounds were speech spectrum noise or a cross-language masker, Chinese listeners benefitted significantly less from spatial separation than did English listeners when a same-language masker was used. These results demonstrate a similar release from energetic masking (noise) due to spatial separation for mono-lingual English and Chinese listeners, but a larger release from informational masking for English than for Chinese listeners. This suggests that one's native language has an effect on informational but not on energetic masking. Moreover, when the competing speech is uninterpretable to the listener, it is indeed easier for a monolingual Chinese listener to segregate the Chinese talker from an English background than it is for a monolingual English listener to segregate an English talker from a Chinese background. It would be interesting to determine what would happen when the target is in one's second language but the masker in one's first language. A secondary point of interest is whether any such effect would be modulated by one's first language (L1-Chinese, L2-English).

Currently, we are testing 24 native Mandarin speakers who were born in China and later immigrated to Canada. 12 participants left China before completing elementary school and 12 after completing it. Each participant is asked to attend two experimental sessions that follow the methodology used by Wu et al. (accepted): one using English target sentences and the other using Chinese target sentences. Unlike in the previous study, where English was uninterpretable to participants, in the current study, English is interpretable to participants—possibly more so for those who immigrated earlier in life than later. Will the extensive exposure to English as a second language affect the degree to which spatial separation

produces a release from masking when the masker is either from one's first language or from one's second language?
Preliminary results and their implication will be discussed.

Acknowledgments

This research was supported by a grant from the Natural Science and Engineering Research Council of Canada (RGPIN-9952). We thank Meiying Qin, Selasie Ametorwo and Carolyn Yeung for their assistance.

References

Ezzatian, P., Avivi, M. & Schneider, B.A. (2010).Do nonnative listeners benefit as much as native listeners from spatial cues that release speech from masking?*Speech Communication*, *52*, 919-929.

Helfer, K. S. (1997). Auditory and auditory-visual perception of clear and conversational speech.*Journal of Speech Language and Hearing Research*, *40(2)*, 432-443.

Wu, X.-H., Yang, Z.-G., Huang, Y., Chen, J., Li, L., Daneman, M., Schneider, B.A. Cross-language differences in informational masking of speech by speech: English versus Mandarin Chinese. *Journal of Speech, Language, and Hearing Research*, accepted.