

**LISTENING TO SEMANTICALLY ANOMALOUS SENTENCES MASKED BY NOISE
AND COMPETING SPEECH IN A SECOND LANGUAGE: A CROSS-LANGUAGE
STUDY ON KOREAN-ENGLISH BILINGUALS.**

Avivi-Reich^{1,2}, M., Bae, M.H^{1,2}., Kang, Y^{3,4}.,& Schneider, B.A.^{1,2}

1. *Department of Psychology, University of Toronto Mississauga, Canada*
2. *Centre for Research on Biological Communication Systems, University of Toronto Mississauga*
3. *Department of Linguistics, University of Toronto, Canada*
4. *Centre for French and Linguistics, University of Toronto Scarborough*

E-mail: meital.avivi@utoronto.ca

Abstract

Wu et al. (2011) examined masking of speech by speech in Mandarin Chinese listeners. The results of their study showed that Chinese listeners benefitted less from spatial separation than did English listeners when a same-language masker was used; implying that the effectiveness of the spatial separation cue with respect to stream segregation is language dependent. In this study we have asked 24 native Korean speakers to perform a similar task and repeat nonsense sentences in Korean played with one of three types of background noises: 1) steady speech-spectrum noise; 2) same-language speech (Korean); and 3) morphologically similar cross-language speech (Chinese). These conditions were played in 4 different SNRs. The 50% correct corresponding SNR thresholds and the slopes of the psychometric functions were calculated and analyzed. The current results show that when listening to Korean the release from masking seems to be also smaller than when listening in English.

Most daily interactions take place in noisy environments rather than in quiet. This is known to create an acoustic challenge for those who attempt to carry out a conversation in such settings, for some more so than for others. One of the most common complaints made by older adults is related to difficulties they encounter when communicating in noise (CHABA, 1988), however, younger adults may also find these acoustic conditions challenging, especially if they are required to perform while using their second language (L2). A growing number of studies demonstrate the difficulties non-native listeners experience when attempting to perceive and understand acoustically disturbed signal (e.g., Ezzatian, Avivi, & Schneider, 2010; Lu-Feng, 2010; Mattys et al., 2010). Their ability to successfully meet the challenges presented by additional background noise is poorer than that of a native listener even when they have been intensely exposed to the second language from an early age (Ezzatian, Avivi and Schneider, 2010). While more information on the effects that a first language (L1) may have on the speech perception of a second language (L2) is accumulating, the possible effects of intense exposure to

a second language (L2) on the speech perception of the first acquired language (L1) remain much less addressed. The question of how an exposure to a second language may affect speech perception performance when listening to the first native language is not less important, since in most cases people do not completely abandon their native language but rather continue to communicate in both. The proportion of people who are using more than one language on a daily basis is increasing because of increases in immigration. In addition, an increasing number of companies and businesses are becoming international in scope, requiring many to communicate in both their first and second language.

From a pragmatic point of view, being able to communicate efficiently in a noisy multi-talker environment is an essential skill. Not being able to efficiently communicate in the presence of other talkers might lead to social isolation and loss of psychological/emotional and other support sources. On a theoretical level, it is important to determine the causes of these difficulties. In nonnative listeners we can consider several possible sources such as a reduced ability to take advantage of context or auditory cues (Ezzatian, Avivi, & Schneider, 2010), limited mastery of the second language's phonetics, slower access to the lexicon, etc. When comparing the results of several studies, which have tried to quantify the contribution of those sources, the complexity of the listening situation seems to be a key variable which must be considered (e.g., Schneider, 2011). In this study we choose to address one acoustic feature which can also be used as a cue when attempting to segregate a target voice from any distracting background sounds, namely spatial separation.

When a target voice is spatially separated from the background noise it allows the listener to take advantage of auditory cues, such as interaural timing differences and SNR differences, to enhance stream segregation. In order to further explore the sources of possible differences in the ability to take advantage of spatial separation, we changed only the apparent location of the target using the precedence effect. If the target voice is played over two loudspeakers located to the left and right of the listener, with the sound on the left loudspeaker slightly lagging behind that on the right, the listener perceives the target voice as emanating from the right. If the masker is played at the same time as the target voice, but with the lag reversed, the masker will be perceived as coming from the left, creating a perceived spatial separation which is achieved without altering the SNRs at each ear. Previous studies, which were done using perceived spatial separation, showed that the ability to use it reduces with age (Murphy et al., 2006) and is somewhat language based (Wu et al., 2010). More specifically, while no difference was found between native and non-native listeners when listening and repeating semantically anomalous sentences in English (Ezzatian, Avivi, & Schneider, 2010), differences were found when monolingual speakers of English and Mandarin Chinese listened in their own language when the target sentence was masked by one of three maskers (same language two-talker speech, cross-language two-talker speech, and speech-spectrum noise; Wu et al., 2010). In the latter study, both groups benefitted equally from spatial separation when the maskers were speech-spectrum noise or cross-language. However, Chinese listeners benefitted less from spatial separation than did English listeners when a same-language masker was used. The results of this cross-language

study implied that the effectiveness of the spatial separation cue with respect to stream segregation is language dependent. In order to further explore the effect of L1 and bilingualism on speech perception and comprehension in noise, the current study was done in Korean which is a morpheme-based language similar to Chinese but lacks tonality. More specifically, we wanted to test whether the extent to which spatial separation can be used as a cue may vary based on the type of masker and the age at which an extensive use of the second language has started.

Method

Participants

Twenty four young native Korean-speakers with normal hearing were recruited from University of Toronto, Mississauga. Twelve of the participants arrived to North America before the age of 14 years and 12 arrived at a later age.

Stimuli and Procedure

The grammatically correct but semantically anomalous English Freyman sentences (Helfer, 1997) were translated into Korean and were recorded using a female native Korean speaker. These sentences contain three keywords corresponding to the subject, verb, and object of the simple declarative structure. The order of these components in the Korean version was *subject + object + verb*. The subject and object words of all sentences consisted of two morphemes, whereas the verb did not follow this rule due to structural differences in Korean versus Chinese. These sentences were organized into 24 lists, with 13 sentences in each list, and were played over loudspeakers in a sound-attenuating booth. These sentences were played along with one of three maskers (steady-state speech spectrum noise, two other female talkers producing semantically-anomalous sentences in Chinese, two different female talkers producing semantically-anomalous sentences in Korean). The spatial location of both the target sentences and the masker was varied using the precedence effect. In the co-located condition, the target and the masker were both presented over the right and left loudspeakers with the left loudspeaker lagging by 4ms behind the right, giving rise to the impression that both the target sentences and the masker were emanating from a position to the right of the listener. In the spatially separated condition, while the target sentences on the left loudspeaker were slightly lagging behind that on the right, the masker was played at the same time as the target voice, but with the lag reversed. As a result the target perceived as coming from the right while the masker was perceived as coming from the left, creating a perceived spatial separation. Each combination of masker type and location was played using 4 different signal to noise ratios (SNR): 6, 1, -4, and -9dB, in order to later assess the SNR at which the participants perceives 50% of the target words correctly. Participants were asked to repeat the sentence they heard and each morpheme in each target word was scored as either correct or incorrect.

Results

The initial results of this study show that similar to what was found in Chinese (Wu et al., 2011), Korean native listeners present smaller release from masking due to spatial separation when listening to Korean, than that which was found in English native speakers. In addition, the cross-language two-talker masker (Chinese) interfered with the perception of Korean more than the English cross-language masker did for native Chinese listeners, and more than the Chinese cross-language masker did for native English listeners in Wu et al. (2011). As can be seen in Fig. 1, no significant difference was found in the extent to which perceived separation was used to enhance release from masking between early and late arrivals. This lack of age of arrival main effect implies that the ability to use this acoustic cue is most likely not dependent on individual linguistic experience.

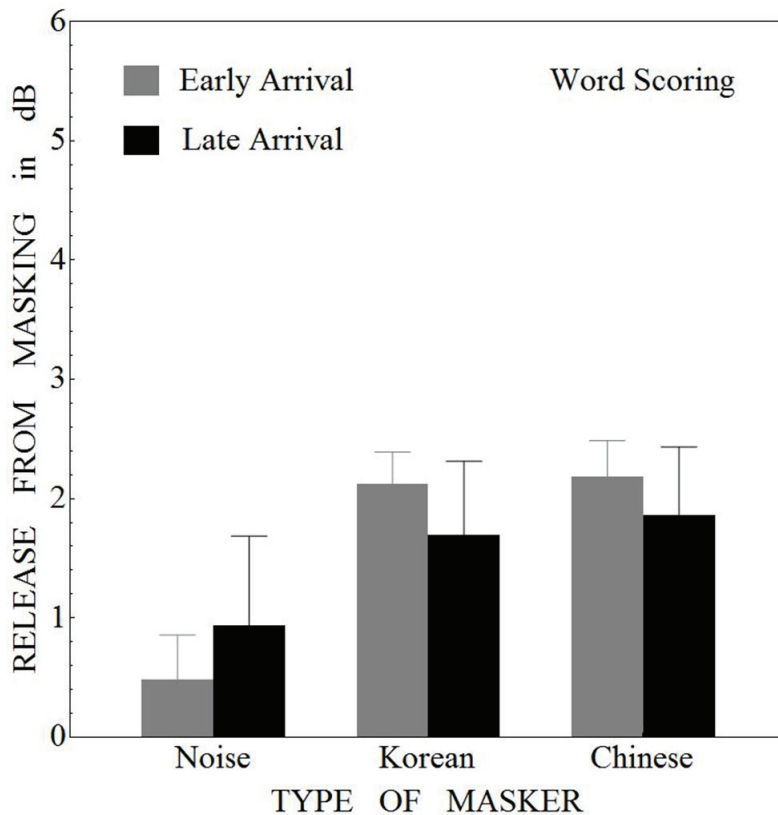


Fig. 1. Mean threshold differences between pairs of maskers for the Conditions in which maskers and targets were co-located versus those where they were, representing release of masking due to spatial separation, separated for Early and Late Arrivals to Canada.

Acknowledgments

This research was supported by a grant from the Natural Science and Engineering Research Council of Canada (RGPIN-9952).

References

- CHABA. (1988). Speech understanding and aging. *The Journal of the Acoustical Society of America*, 83, 859 – 895.
- 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. *J. Sp. Lan. Hear. Res.*, 40, 432-443.
- Lu-Feng, S. (2010). Perception of acoustically degraded sentences in bilingual listeners who differ in age of English acquisition. *Journal of speech, language, and hearing research* : 53 (4), 821 – 835.
- Mattys, S.L., Carroll, L.M., Li, C.K.W & Chan, S.L.Y.(2010). Effects of energetic and informational masking on speech segmentation by native and non-native speakers. *Speech Communication*, 52(11), 887 – 899.
- Schneider, B.A. (2011). How age affects auditory-cognitive interactions in speech comprehension. *Audiology Research*, 1(1S), 34-39.
- Wu, X., Yang, Z., Huang, Y., Chen, J., Li, L., Daneman, M., & Schneider, B. (2011). Cross-language differences in informational masking of speech by speech: English versus Mandarin Chinese. *Journal of Speech Language and Hearing Research*, 54(6), 1506-1524.