Influence of native and non-native multitalker babble on speech recognition in noise

Chandni Jain, Sreeraj Konadath, Bharathi M. Vimal, Vidhya Suresh

Department of Audiology, All India Institute of Speech and Hearing, Manasagangothri, Mysore, India

Abstract

The aim of the study was to assess speech recognition in noise using multitalker babble of native and non-native language at two different signal to noise ratios. The speech recognition in noise was assessed on 60 participants (18 to 30 years) with normal hearing sensitivity, having Malayalam and Kannada as their native language. For this purpose, 6 and 10 multitalker babble were generated in Kannada and Malayalam language. Speech recognition was assessed for native listeners of both the languages in the presence of native and non-native multitalker babble. Results showed that the speech recognition in noise was significantly higher for 0 dB signal to noise ratio (SNR) compared to -3 dB SNR for both the languages. Performance of Kannada Listeners was significantly higher in the presence of native (Kannada) babble compared to non-native babble (Malayalam). However, this was not same with the Malayalam listeners wherein they performed equally well with native (Malayalam) as well as non-native babble (Kannada). The results of the present study highlight the importance of using native multitalker babble for Kannada listeners in lieu of non-native babble and, considering the importance of each SNR for estimating speech recognition in noise scores. Further research is needed to assess speech recognition in Malayalam listeners in the presence of other non-native backgrounds of various types.

Introduction

Perception of speech is the ability of the listener to perceive the acoustic waveform produced by a speaker as a string of meaningful words and ideas and attend to it. Comprehensive understanding of speech depends on various noise conditions like, signal to noise ratio (SNR) and types of background noise. Speech recognition in noise is clinically used to assess the auditory functions in individuals with hearing sensitivity within normal limits and with hearing loss.

Speech recognition in noise is affected by two types of processes, namely signal driven and knowledge driven. The main purpose of signal driven processes is to analyze different acoustic sources of the signal to noise and types of background noise. Speech recognition in noise is clinically used to assess the auditory functions in individuals with hearing sensitivity within normal limits and with hearing loss.

Speech recognition in noise is affected by two types of processes, namely signal driven and knowledge driven. The main purpose of signal driven processes is to analyze different acoustic sources of the signal to noise and types of background noise. Speech recognition in noise is clinically used to assess the auditory functions in individuals with hearing sensitivity within normal limits and with hearing loss.

Speech recognition in noise is affected by two types of processes, namely signal driven and knowledge driven. The main purpose of signal driven processes is to analyze different acoustic sources of the signal to noise and types of background noise. Speech recognition in noise is clinically used to assess the auditory functions in individuals with hearing sensitivity within normal limits and with hearing loss.

The knowledge driven process in speech recognition in noise can be studied by comparing the listeners with different native languages in the presence of different multitalker babble i.e., the multitalker babble of native as well as foreign languages. The speech perception in the presence of different competing noise would result in masking of the target speech. This could be due to energetic masking and informational masking. Energetic masking occurs as a result from competition between target and masker at the periphery of the auditory system and informational masking refers to everything that reduces intelligibility of speech when energetic masking is accounted for.

In the past, a number of studies have been done to study the effects of babble on the listener’s recognition for sentences in their native language. Studies have shown that the linguistic content of the masking speech signal can influence the speech recognition. It is reported that the performance of the listener degrades when the competing speech is spoken in the listeners’ native language versus a language that is unfamiliar to them.

In another study Calandrucio et al. investigated the effect of linguistically and phonetically close and distant maskers on speech recognition. They reported that the speech recognition performance increased as the target-to-masker linguistic distance was increased. In a study by Russo and Pichora-Fuller they reported that the word recognition performance was poorer in the presence of multitalker babble compared to familiar music.

The objective of this research was to investigate the effect of native and non-native babble backgrounds on speech recognition performance. India is a multilingual country with twenty-three constitutional-
Materials and Methods

Participants

Sixty participants in the age range of 18 to 30 years with normal hearing sensitivity participated in the study. Thirty native speakers of Malayalam and thirty native speakers of Kannada were selected for the study. Participants in both the groups did not know how to speak, read and write the non-native language and they had English as their second language. The listeners who participated had normal hearing, as indicated by their four-frequency (500 Hz, 1000 Hz, 2000 Hz, 4000 Hz) pure-tone average threshold of ≤15 dBHL, ‘A’ Type tympanogram with acoustic reflex thresholds in normal limits. A structured interview of these listeners was taken in their native language to make sure that none of them had any difficulty in understanding speech in daily listening conditions. It was also ensured that they did not have any history of neurologic or otologic disorder. All the listeners were native speakers of Kannada/Malayalam and their participation was voluntary. They were not paid for their participation in the study.

Materials

For the aim of perception experiment in Kannada and Malayalam languages phonetically balanced words in Malayalam language was developed and the already existing Kannada wordlist[12] was used. For the purpose of development of words for speech recognition in Malayalam language, 100 monosyllabic words were collected and they were distributed into four sub lists with 25 words each. The selected words were all verbs and all the four sub lists were made phonetically balanced. Phonetic balancing was done based on the reference frequency distribution of Malayalam language. The selected 100 words were recorded digitally in a sound-proof booth, using a computerized speech lab at a sampling rate of 44.1 kHz with 24-bit resolution. A microphone was placed at a distance of 20 cm from the speaker’s mouth and the speaker was informed to articulate all words clearly. These words were recorded by an adult female who was a native speaker of Malayalam. The speaker used for recording was a bilingual speaker, with Malayalam as her first language and English as a second language. She was instructed to read the words in the list as natural as possible. The individual wave files of the words were stored onto a computer hard disk and were normalized by using Adobe Audition (Version 3.0) software.

To ensure that the developed lists were of equal intelligibility, the speech recognition scores were found out on 30 native Malayalam speakers and the results showed that all the lists were of equal difficulty. Speech recognition scores of all the participants ranged from 95 to 100% across the lists. Word list in Kannada was adapted from Sreela and Devi.[12] This list also consisted of 100 words distributed in four sub lists with 25 words each. The Kannada wordlist used was also phonetically balanced and consisted of verbs and the intelligibility of this list was checked by the authors on 30 native Kannada speakers.

Generation of multitalker babble

For the purpose of assessing speech recognition in the presence of noise, multitalker babble was generated in Kannada and Malayalam languages. For the construction of babble, speech was recorded from 20 individuals using different standardized passages in Kannada and Malayalam. The passages consisted of 304 words and 307 words respectively in Kannada and Malayalam.[14] Ten native speakers of Malayalam and ten native speakers of Kannada language with equal number of males and females were selected to generate ten talker and six talker babble. Same talkers were used for the generation of ten talker and six talker babble and all the talkers were native speakers of the language with English as their second language. Further, Kannada multitalker babble was mixed with Kannada and Malayalam words at 0dB SNR and -3dB SNR with the use of MATLAB code. Similarly, Malayalam multitalker babble was mixed with Malayalam and Kannada words to meet the objectives of the study. The onset of multitalker babble preceded the onset of the word by 600 ms and continued till 600 ms after the end of each word in the recorded list.

Procedure for speech recognition testing

Speech recognition was assessed using the above described test materials on 60 participants (18 to 30 years) with normal hearing sensitivity, having Malayalam and Kannada as a native language (30 in each language group). Participants in both the groups did not know how to speak, read and write the non-native language.

The testing was carried out at 80 dBSPL, wherein the stimulus was delivered through a computer connected to a calibrated double channel diagnostic audiometer (OB - 922). The subjects were instructed to listen to the stimulus carefully through headphones (TDH 39) and to repeat the words that he/she hears. Speech recognition of the native Malayalam speaker was estimated using Malayalam word list in the presence of Malayalam and Kannada babble. An assessment was done using ten and six talkers babble at 0 and -3 dB SNR. Randomization was done to ensure that the order of presentation of lists, SNRs and multitalker babble were not in the same sequence for all the subjects, to avoid the practice effect. A similar procedure was followed with respect to Kannada speaker wherein Kannada listeners heard Kannada words in the presence of Kannada and Malayalam babble. The words correctly repeated were given score ‘1’ and wrongly repeated were given ‘0’ with a maximum possible score of each list as 25. The responses were scored during the session by the tester.

Results and Discussion

Statistical analysis was performed using SPSS 20 software.
Comparison of speech recognition of Kannada and Malayalam speakers in the presence of native vs. non-native babble, various SNRs and number of talkers in multitalker babble were done. Descriptive statistics to estimate the mean and standard deviation and mixed ANOVA was done where SNRs and multitalker babble were compared for within subject variable and language of the subject was considered as between subject variable.

Results of speech recognition scores across SNRs (0 and -3 dB) using ANOVA showed that SNR had significant main effect \( [F (7, 833) =91.66, P<0.01] \) for both Kannada and Malayalam listeners for six and ten multitalker babble (Figures 1 and 2). A further pair wise comparison was done using Boneferroni test which revealed significant differences between SNRs for both 10 and 6 talkers babble. In Figures 1 and 2, x-axis represents different SNRs for six talkers and ten talker babble and the y axis represents average speech recognition scores and standard deviation (SD) for Kannada and Malayalam speakers. This result is reported by combining the scores obtained in the presence of native and non-native babble for both the listening groups. In Figure 1, red bar indicates Kannada speakers hearing Kannada words (averaged across native and non-native babble types) and blue bar indicates Malayalam speakers hearing Malayalam words (averaged across native and non-native babble types). It is evident from Figure 2 that scores were better for 0 dB SNR compared to -3 dB SNR. The result of the present study is in agreement with the studies done in the past.\textsuperscript{15-18} These studies have reported that speech recognition becomes poorer with an increase in the level of the noise with respect to the speech signal.

Results of speech recognition scores to compare the effect of number of talkers in the multitalker babble (10 and 6 talkers babble) using ANOVA showed significant main effect \( [F (7, 833) =91.66, P<0.01] \) for both Kannada and Malayalam listeners. Further pair-wise comparison was done using Boneferroni's test which revealed significant difference in speech recognition scores across six and ten talker babble for Kannada speakers and there was no significant difference across Malayalam speakers for both 0 and -3 dB SNR. In Figures 3 and 4 speech recognition of Kannada and Malayalam listeners is compared in the presence of six and ten talker babble. In Figures 3 and 4, red bar indicate Kannada listeners hearing Kannada words (averaged across native and non-native babble types) and blue bar indicates Malayalam listeners hearing Malayalam words (averaged across native and non-native babble types). This result is in concurrence with the studies done in the past.\textsuperscript{15,16,18-20} The reasoning could be that perceptibility of speech in multitalker babble decreases as the number of talkers in the noise increases. However in Malayalam listeners no difference in speech recognition was observed for six and ten talker babble.

Comparison of speech recognition of Kannada and Malayalam listeners was also done in the presence of native vs. non-native multitalker babble using ANOVA which showed a significant main effect \( [F (15, 885) =34.12, P<0.01] \). Further, pair-wise comparison was done using Boneferroni’s test which revealed that for Malayalam listeners there was no significant difference (P>0.05) in speech recognition scores in

![Figure 1. Mean scores and standard deviation across different signal to noise ratio (SNR) for 6 talker babble for Kannada and Malayalam speakers.](image1)

![Figure 2. Mean scores and standard deviation across different signal to noise ratio (SNR) for 10 talker babble for Kannada and Malayalam speakers.](image2)

![Figure 3. Speech scores and standard deviation at 0dB signal to noise ratio (SNR) for 6 (6TB) and 10 talker babble (10TB) in Kannada and Malayalam speakers.](image3)

![Figure 4. Speech scores and standard deviation at -3dB signal to noise ratio (SNR) for 6 and 10 talker babble in Kannada and Malayalam speakers.](image4)
the presence of native vs. non-native babble as shown in Figure 5. In Figure 5, red bar indicates speech recognition scores of Malayalam native speakers using Malayalam words in the presence of non-native (Kannada) babble and blue bar indicates speech recognition scores of Malayalam native speakers in the presence of native (Malayalam) babble. However, in Kannada listeners there was a significant difference (P<0.001) in speech recognition scores in the presence of native vs non-native babble as shown in Figure 6. In Figure 6, red bar indicates speech recognition scores of Kannada native speakers using Kannada words in the presence of native (Kannada) babble and blue bar indicates speech recognition scores of Kannada native speakers in the presence of non-native (Malayalam) babble. Further, to see the interaction of SNR and babble on the two groups (Malayalam listeners vs Kannada listeners) mixed ANOVA was done. Results revealed that there was no significant interaction of SNR and babble in the above mentioned groups [F (1,118)=1.542, P>0.05].

Speech recognition performance of Kannada listeners was better in the background of native (Kannada) babble compared to non-native babble (Malayalam). The results of the present study is in agreement with the study9 done by Russo and Pichora-Fuller where speech identification was found to be better in the familiar background compared to unfamiliar background, but however in the above study the familiar background which was used was music. However, few studies have shown that the scores are poorer using native babble as compared to non-native babble.21 This could be due to that non-native listeners are more adversely affected than native listeners due to energetic and informational masking.22 Studies have also shown that the language of interfering noise can affect the intelligibility of the target speech.20,21

The reason for the result of the present study being not in agreement with studies mentioned above could be attributed to the similarities between two languages being studied as they are of the same origin (Dravidian languages). The familiarity with the non-native language or the acoustic-phonetic similarity of the two languages would have made this difference. Moreover, cognitive factor would play a role where participants would get more distracted with a new language in the background compared to the known language.

However, the reason as to why there was no significant difference in speech recognition scores of Malayalam listeners in the presence of native (Malayalam) and non-native (Kannada) babble is not clear and also there was no significant difference in speech recognition scores with the number of babble. To author's knowledge no other study has compared speech recognition in Malayalam language with different backgrounds. Thus it can be concluded that Malayalam listeners are able to do equally well in the presence of native as well as nonnative babble.

Conclusions

The results of the present study showed better speech recognition scores for native listeners of Kannada compared to non-native listeners (Malayalam) for Kannada multtalker babble. It was also noted that the speech recognition in noise scores was relatively better when there is more number of talkers in multitalker babble for Kannada listeners. When different SNR was compared it was observed that speech recognition scores were better at higher SNR (0dBSNR) compared to lower SNR (-3dBSNR).

The results of the present study highlight the importance of using native multitalker babble for Kannada listeners in lieu of non-native babble and considering the importance of each SNR for estimating speech recognition scores. However, the results of speech recognition scores for Malayalam listeners in the presence of native and non-native babble are not conclusive. Hence, further research addressing the same may be conducted with other non-native languages which would highlight about the recognition abilities of Malayalam listeners.

References

1. Goldinger SD, Pisoni DB, Luce PA. Speech perception and spoken word recognition: research and theory. In: Lass NJ, ed. Principles of experimental phonetics. St. Louis, MO: Mosby; 1996. pp 277-327.
2. Bregman AS. Auditory scene analysis. Cambridge: MIT; 1990.
3. Remez RE, Rubin PE, Berns SM, et al. On the perceptual organization of speech, Psychol Rev 1994;101:129-56.
4. Durlach N. Auditory masking: need for improved conceptual structure. J Acoust Soc Am 2006;120:1787-90.
5. Brouwer S, Van Engen KJ, Calandruccio L, Bradlow AR. Linguistic contributions to speech-on-speech masking for native and non-native listeners: Language familiarity and semantic content. J Acoust Soc Am 2012;131:1449-64.
6. Calandruccio L, Van Engen KJ, Dhar S, Bradlow AR. The effects of clear speech as a masker. J Speech Lang Hear Res 2010;53:1-14.
7. Engen VK, Bradlow AR. Sentence recognition in native- and foreign-language multitalker background noise. J Acoust Soc Am 2007;121:519-26.
8. Calandruccio L, Brouwer S, Van Engen KJ, et al. Masking release due to linguistic and phonetic dissimilarity between the target and masker speech. Am J Audiol 2013;22:157-64.
9. Russo F, Pichora-Fuller MK. Tune in or tune out: Age related differ-
ences in listening when speech is in the foreground and music is in the background. Ear Hear 2008;29:746-60.
10. Mallikarjun B. Fifty years of language planning for modern Hindi – The official language of India. Language India 2004;4:11.
11. Mallikarajun B. Evolution of language policy for education in Karnataka. In: K.S. Rajashree and Udaya Narayana Singh, eds. Papers in applied linguistics I. Mysore: Central Institute of Indian Languages; 2001. pp 122-145.
12. Sreela PK, Devi N. Effect of reverberation on speech identification using hearing aids. Articles based on dissertation at AIISH 2009;8:221-30.
13. Ramakrishna BS, Nair KK, Chiplunkar VN, et al. Some aspect of the relative efficiencies of Indian languages. Bangalore: Indian Institute of Science; 1962.
14. Savithri SR, Jayaram M. Girijamma. Rate of reading in Kannada speakers. J Acoust Soc India 2004;Sec C:1-6.
15. Bronkhorst AW, Plomp R. Effect of multiple speech like maskers on binaural speech recognition in normal and impaired hearing, J Acoust Soc Am 1992;92:3132-9.
16. Bronkhorst A. The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. Acoustica 2000;86:117-28.
17. Rhebergen KS, Versfeld NJ. A speech intelligibility index based approach to predict the speech perception threshold for sentences in fluctuating noise for normal-hearing listeners. J Acoust Soc Am 2005;117:2181-92.
18. Simpson SA, Cooke M. Consonant identification in N-talker babble is a nonmonotonic function of N. J Acoust Soc Am 2005;118:2775-8.
19. Brungart D, Simpson B, Ericson M, Scott K. Informational and energetic masking effects in the perception of multiple simultaneous talkers. J Acoust Soc Am 2001;110:2527-38.
20. Rhebergen KS, Versfeld NJ, Dreschler WA. Release from informational masking by time reversal of native and non-native interfering speech. J Acoust Soc Am 2005;118:1274-7.
21. Lecumberri GML, Cooke MP. Effect of masker type on native and non-native consonant perception in noise. J Acoust Soc Am 2006;119:2445-54.
22. Engen VK, Bradlow AR. Sentence recognition in native- and foreign-language multitalker background noise. J Acoust Soc Am 2007;121:519-26.