Speech-in-speech recognition: effect of language uncertainty

Preeta Singh[1] & Geetha C[2]

Abstract

Speech-in-speech recognition is a common phenomenon, in a multilingual country like India. This phenomenon co- exists with unpredictable change in the background language, especially with English language. The present study aims to study the influence of native vs. non-native language babble, language uncertainty of the babble and role of signal-to-noise ratio on the Kannada sentence recognition. Forty Kannada-English Bilinguals were subjected to Kannada sentence recognition task in the presence of two-talker Kannada babble, two-talker English babble, two- talker mixed (Kannada + English) babble and two-talker interleaved (unpredictable change in background language) at +3 dB SNR, 0 dB SNR and -3 dB SNR. The level of presentation of target sentence was 75 dB SPL. Comparison of the recognition scores across four babble conditions revealed significantly better scores in the presence of two-talker English babble when compared to two-talker Kannada babble, implying the release of masking in the presence of non-native language babble. Further, the presence of language uncertainty in the two- talker interleaved condition reduced the scores when compared to two-talker mixed condition. The influence of other variables like F_0 and LTASS differences influencing the speech-in-speech recognition was negligible and the release of masking can be attributed to the linguistic mismatch of the target and the babble. In addition, SNR was found to be a factor influencing speech-in-speech recognition.

Key words: *language uncertainity, informational masksing,two talker babble, speech in speech recognition*

Introduction

Speech recognition in the presence of multiple language background and multiple talkers is a common phenomenon. The impact of such noise on the speech intelligibility depends on the acoustic characteristic of the noise, semantic characteristics, meaningfulness, linguistic and phonetic content, native or non native language, number of background talkers, gender of the speaker and language uncertainty (Dillon, 1983; Freyman, Helfer, & Balakrishnan, 2007; Kalikow, Stevens, & Elliott, 1977; Van Engen, 2010) among others.

Research studies have tried to examine the effect of these factors on speech recognition scores (SRS) in presence of multi-talker babble (MTB). This task is also known as speech-in-speech recognition. The use of MTB that contains linguistic information would compete with the recognition of the target rather than cause spectral or temporal masking, in other words, speech recognition in the presence of MTB causes less energetic masking (EM) and more informational masking (IM). Studies have shown that the amount of IM can differ when the relationship between the target and masker stimuli is varied (Brouwer & Bradlow, 2014; Brouwer, Engen, Calandruccio, & Bradlow, 2012; Carhart, Tillman, & Greetis, 1969; Engen, 2012).

Several studies have been done to assess the effect of linguistic and phonetic similarity between the languages and listeners familiarity of the language in speech-in-speech recognition task. These studies have shown that unfamiliar language masker leads to more masking release compared to native language or a familiar language (Brouwer et al., 2012; Jain, Konadath, Vimal, & Suresh, 2014). Further, decreasing similarity between the target and masker has been found to decrease the SRS. For example, the English language sentence recognition in two-talker babble of Mandarin, resulted in better SRS as Mandarin is dissimilar to English, when compared to that of English babble (Calandruccio, Brouwer, Van Engen, Dhar, & Bradlow, 2013).

Unlike the above findings, studies on speech-in- speech recognition of Indian languages have shown varied results. Anitha (2003) studied the effect of ten-talker babble of different languages (Kannada, Hindi, Malayalam) on SRS of Kannada phonetically balanced words in 40 Kannada speakers. SRS in the presence of ten- talker babble was poorer when compared to SRS in speech noise, whereas no significant difference was obtained across different language babbles. The authors' opinion about the result was that the major factor for masking was the spectrum of the babble not the linguistic or semantic content in the babble.

However, in the above study, the number of talkers in the babble was too large to assess the effect of linguistic context. Nonetheless, results of the study done by Jain et al., (2014) showed that the Kannada speakers performed better in the presence of Kannada MTB than non-native Malayalam MTB. The authors speculate the reason for the poorer performance by Kannada group in the presence of non-native MTB as attention based factors, where the participants are distracted by non-

1. preeta91@gmail.com
2. geethamysore.cs@gmail.com

native language, leading to poorer scores in presence of non native MTB.

The above mentioned studies studied the effect of different Indian languages. Research on the influence of a non-Indian language babble on recognition of Indian language is sparse. A study done by Vineetha, Suma, and Nair (2013) reported no difference in the SRS obtained by Kannada-English bilinguals in the perception of Kannada words in the presence of four-talker Kannada and English babble at 5, 0,-5,-10 and -20 dB SNR.

Various factors could have influenced the difference obtained among studies. The factors could be difference in type of test material (words, sentences), number of takers in the MTB, language familiarity and similarity between the target language and the background language.

Another factor which could affect the speech in speech masking is uncertainty of the language of the babble. Various studies using non-speech stimuli have studied the effect of uncertainty of the masker tone. They reported that the uncertainty adversely affects thresholds of tonal target (Durlach, Mason, Shinn-Cunningham, et al., 2003; Kidd, Mason, & Arbogast, 2002; Neff, Dethlefs, & Jesteadt, 1993; Watson, Kelly, & Wroton, 1976). Similarly, uncertainty can also be induced by changing the language of babble. The study done by Brouwer and Bradlow (2014) investigated the effect of variation in the target-background language relationship (contextual variation) on the speech-in-speech recognition on 48 native English speakers. Two experiments were carried out to test the SRS of English sentences. The first experiment had English sentence recognition in the presence of two-talker English babble, two-talker Dutch babble and interleaved language condition (interleaved English and Dutch babble). In the interleaved condition, the English and Dutch Babble were interleaved such that it switched languages 30% of the time. They reported poorer scores in the condition where babble was a mixture of both the languages and this was due to IM.

On the contrary, studies which have induced uncertainty in terms of spatial location of masker or number of talkers and gender of the talker of babble (Brungart & Simpson, 2004; Freyman et al., 2007; Jones & Litovsky, 2008) reported not much of an effect of uncertainty on the SRS.

The above experiments studying IM have been studied in various SNRs. They found that SNR does affect the amount of IM and EM. EM increases with decreasing SNR (Calandruccio, Dhar, & Bradlow, 2010), whereas, IM was found to influence speech recognition the most when linguistic content of both target and babble are audible, and to influence the least at easy SNR and difficult SNR where the target or the babble is least audible (Engen & Bradlow, 2007; Engen, 2010).

Due to the advent of globalization, the phenomenon of bilingualism/multilingualism has become common. In a multi-lingual country like India, English has coexisted with the indigenous languages since the British rule, and now it is one of the official languages of the country. English is used as a language of education in most schools, colleges and universities. In the domestic setting, along with English, other local language is also used and spoken. Hence, in the classrooms or in different social settings, the mixture and switching of codes have become inevitable (Harini & Shyamala, 2008).

The earlier studies on Indian languages have studied only the influence of an Indian language on another Indian language. These studies have used MTB either with 4-, 5-, 6- or 10-talker babble (Anitha, 2003; Jain, et al, , 2014). In Anitha's study, the results revealed no significant difference in terms of language of the masker. This could be because of the number of speakers used in the babble. When the number of talkers is more, the background signal resembles the speech noise and hence, the semantics is lost. Therefore, there is more influence of EM than IM (Carhart, Johnson, & Goodman, 1975). In a study done by Van Egan and Bradlow (2007), significant difference was observed for two-talker babble than six-talker babble. Similarly, Freyman, Balakrishnan, and Helfer (2004) compared two- talker babble versus 3-, 4-, 6- and 10-talker babble and reported that maximal informational masking occurs in two-talker babble background. Hence, it is important to study the effect of speech babble with lesser number of talkers in order to test the influence of language content of the masker.

In addition, there is an evidence that the performance decreases if the language of the babble changes unpredictably. In Indian context, there is a lot of code mixing and code switching, especially with English words (Harini & Shyamala, 2008; Mathew, 2012) and hence, the unpredictable change in the background language is a common phenomenon. Hence, it is important to study the effect of English language background on speech recognition and the effect of unpredictable changes in the language. Further, the role of SNR cannot be neglected as it also plays an important role in influencing the amount of EM and IM in speech-in-speech recognition task (Calandruccio et al., 2010; Engen & Bradlow, 2007; Van Engen, 2010).

The aim of the present study was to study the effect of native and non-native language babble on SRS of Kannada sentences, to study the effect of language uncertainty on the SRS of Kannada sentences, and to study the role of SNR on speech-in-speech recognition in Kannada-English bilinguals

The objectives of the study were:

1. To obtain SRS of Kannada sentences in the presence of two-talker Kannada babble, two-talker

English babble, two-talker mixed (Kannada+ English) babble and two-talker interleaved (background language switching between Kannada and English unpredictably) babble at +3 dB SNR, 0 dB SNR and -3 dB SNR.

2. To compare the SRS in presence of two-talker Kannada babble and two-talker English babble.

3. To determine the effect of language uncertainty by comparing the SRS in the presence of two-talker interleaved babble with other two-talker babble conditions.

4. To determine the effect of SNR on the SRS across the four two-talker babble conditions.

Method

Participants: A group of 40 participants in the age range of 18-30 years (mean = 24.9; SD = 3.153; 20 males and 20 females) were selected based on the following criteria..

1. All the listeners were native speakers of Kannada with English being the second language exposed at least from 5th grade.

2. All the participants had a minimum of Xth grade education in English medium School and Kannada as the second language.

3. The participants were considered as Kannada-English bilinguals if they obtained a score of two or above in English in the International second language proficiency rating scales (ISLPR) developed by Wylie (2006).

4. They had hearing sensitivity less than or equal to a four frequency puretone average (at 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz) of 15 dB HL (WHO, 2008).They had either 'A' or 'As' type of tympanogram with ipsi and contralateral reflexes present within 100 dB HL (Jerger, 1970).

5. They had SPIN scores of 60% or above at 0 dB SNR.They had presence of oto-acoustic emissions in both ears.They reported of no other otological or neurological symptoms, and no other speech or language problems.

Instrumentation: Routine audiological evaluation was carried out using a calibrated two channel diagnostic audiometer (GSI-61) for puretone threshold estimation and speech audiometry. TDH-39 headphones coupled with MX/41AR ear cushions and Radio ear B-71 bone vibrator were the transducers used. A calibrated GSI-tympstar (Grason-Stadler Incorporation, USA) clinical immittance meter was used for tympanometry and acoustic reflex measurement. For the experimental task, a HP Notebook with software Adobe Audition v3 was used for recording, signal generation, processing and mixing. To record the Kannada and English passage for the construction of the babble, MOTU Microbook II, an audio interface was used.

*Material used:*The speech recognition thresholds were obtained using Kannada paired words developed at the Department of Audiology. The Kannada Phonemically Balanced word list developed by (Yathiraj & Vijayalakshmi, 2005) was used to find SRS for routine audiological evaluation and SPIN testing. Kannada sentence list developed by Geetha, Kumar, Manjula, and Pavan (2014) was used to find SRS for sentences in quiet and in presence of two-talker babble. This test consists of twenty five homogeneous lists with ten sentences under each list. For the construction of the babble, Kannada sentences from standardized passage of 300 words in Kannada (developed by Savithri and Jayaram (2005)) and English sentences from the standardized English rainbow passage (Fairbanks, 1960). Were used.

Test environment: The test was carried out in an air conditioned sound treated double room suite with ambient noise levels within permissible limits ANSI S3.6 (1999).

Procedure: The test procedure was carried out in two different phases:

Phase 1: Development of two-talker English babble, Kannada babble, interleaved babble and mixed babble

Phase 2: Measurement of SRS in the presence of four two- talker babble conditions

a) Phase 1: Development of two-talker Kannada babble, English babble, interleaved babble and Mixed babble in 3 different SNR condition*:*

For the construction of the four two-talker babble, Kannada sentences from standardized passage of 300 words in Kannada developed by Savithri and Jayaram (2005) were recorded by two native speakers (one male and one female). The recording microphone was placed 20 cm in front of the mouth of the speaker and the speaker was asked to articulate the words clearly. The passage was recorded digitally in a sound proof booth using MOTU Microbook II and mixed using Adobe Audition version 3 at a sampling rate of 44.1 kHz with 24-bit resolution. The same procedure was used to record English sentences using the standardized English rainbow passage (Fairbanks, 1960).

The sentences recorded in Kannada by each talker were concatenated, with no silent intervals between the sentences to create a string of sentence. The order of concatenation was varied such that the sentences spoken by male and female talker were not the same at a point of time. This was also done for English sentences spoken by one male and one female. For each of the four two-talker babble conditions and three SNRs (4*3), twelve strings of sentences of each talker (one male and one female native Kannada speakers, and one male

and one female non- native English speaker) were created. The sentences were normalized before and after concatenation. The above procedure was done using Adobe audition version 3.

Kannada sentences spoken by two speakers' one male and one female were mixed to create two-talker Kannada babble. Similarly, two-talker English babble was constructed. For two-talker Mixed babble, English sentences spoken by one speaker and Kannada sentences spoken by another speaker of the opposite gender was mixed. Whereas the two-talker interleaved babble was constructed by interleaving (changing) the two-talker Kannada babble and two-talker English babble.

Construction of two-talker Kannada babble, English babble, mixed babble and interleaved babble

The two-talker Kannada babble was constructed by mixing two strings of Kannada sentences spoken by one male and one female talker using Adobe audition version 3. Out of the twelve strings of Kannada sentences spoken by two talkers, three strings of sentences of each talker were used to create three sets of two-talker Kannada babble for the three SNR condition. Similarly three strings of English sentences of each talker (one male and one female) were mixed to create three sets of two-talker English babble.

The two-talker mixed babble was constructed by mixing one string of Kannada sentences and one string of English sentences such that both are spoken by opposite gender and not the same gender. For example, one string of Kannada sentence spoken by male talker was mixed with one string of English sentence spoken by female talker. For the three SNR condition, three sets of two-talker mixed babble were created.

In order to construct two-talker interleaved babble, three sets of the two-talker English and two-talker Kannada babble constructed in section 3.5.1.1. were used. One set of two-talker Kannada and one set of two-talker English babble were interleaved such that it changes from Kannada to English or vice versa 30% of the time. The duration of interleaving condition was altered with respect to duration of the target sentences. For example, for the first two target sentences, the background language was two-talker Kannada babble and for the next three target sentences, the background changed to two-talker English babble. The next three sentences had two-talker Kannada babble and the last two sentences had two-talker English babble. Hence, for one set of interleaved condition, the language of the babble changed thrice for ten target sentences and the pattern of language change or uncertainty was different for each of the SNR condition.

The two-talker Kannada babble, English babble, mixed babble and interleaved babble were mixed with the target standardized Kannada sentences spoken by a female speaker. Each set of two-talker babble was mixed with ten target sentences such that the onset of babble precedes the onset of word by 1second and continues till 1 second even after the end of the last sentence. The inter target interval of 6 second was given in order to give sufficient time for the oral response. The level of the target sentences was fixed at 75 dB SPL and level of the babble tracks were varied to produce a target to babble ratio of +3 dB SNR, 0 dB SNR and -3 dB SNR.

Hence three sets of each two-talker Kannada babble, English babble, mixed babble and interleaved babble were mixed with the different set of target sentences were created to test SRS of Kannada sentences in these SNR.

The sentences were not repeated for any of the conditions to avoid practice effect, and the order of presentation of each condition and the SNR were randomized to eliminate the order effect.

Finally, the three sets of each two-talker Kannada babble, English babble, mixed babble and interleaved babble with target sentences were constructed in three SNR and were presented to the listeners as illustrated in Figure 1.

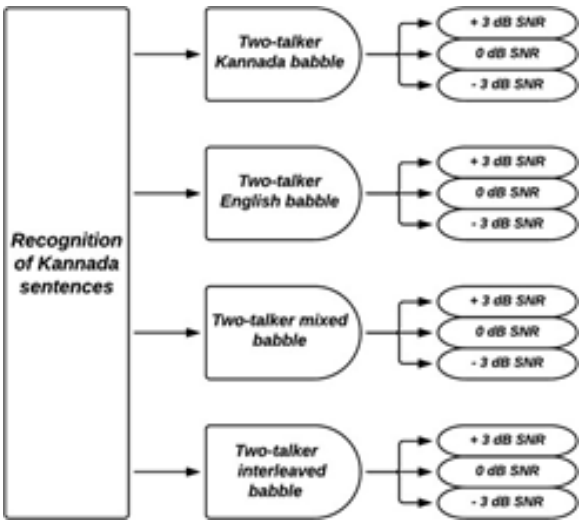

Figure 1: Illustration of the test conditions.

b) Phase 2: Measurement of SRS in the presence of four two-talker babble conditions

The participants were seated in a comfortable chair and the sentences were presented through HP notebook and calibrated i-Ball circum aural headphones diotically. All the participants were instructed to repeat the target sentences orally. The examiner scored the words identified correctly in the score sheet containing the target sentences. Every sentence in the sentences list had four key words and scoring was based on the correct identification of the key words in the sentence. The maximum number of keywords for each SNR condition was 40.

Results

The objectives of the present study were to develop two-talker Kannada babble, two-talker English babble, two-talker mixed babble and two-talker interleaved babble, and to compare the SRS of Kannada sentences in Kannada-English bilinguals in the presence of the above babbles in three SNRs.

1. Development of four different speech babbles:

Four different speech babbles were developed, that is, Kannada babble, English babble, mixed babble and interleaved babble. All the babble had one male and one female talker. As the LTASS and the F0 of the babble in comparison to that of the target stimuli has been reported to be factors affecting the speech recognition, the descriptive analysis of LTASS and F0 of the babble used in the present study are given below. The LTASS of the target Kannada sentences and four two-talker babbles are shown in Figure 2.

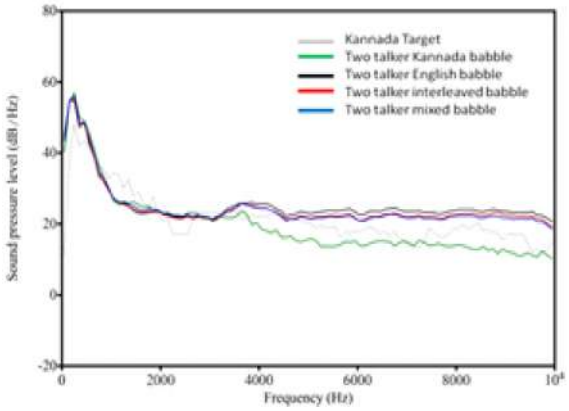

Figure 2: LTASS of different babbles and target sentence.

As the figure depicts, the LTASS of two-talker English, mixed and interleaved babble were similar and when the LTASS of these were compared to the target sentences, the target had lesser energy at high frequencies and equal energy at low frequencies. However, the LTASS of target stimuli was comparable to the Kannada babble at all frequencies.

The F0 of the target stimuli (spoken by a female) and F0 of two-talker babble in Kannada (male and female) and English (one male and one female) were 225 Hz, 145 Hz, 210 Hz, 130 Hz and 205 Hz, respectively. The difference between the F0 of the target Kannada sentences and the male speaker of the Kannada babble was 80 Hz and female speaker of the Kannada babble was 15 Hz, and the difference with male speaker of the English babble was 95 Hz and female speaker of the English babble was 20 Hz.

2. Measurement of SRS in the presence of different babbles

The scores of sentence recognition obtained for each of the three SNRs under four two-talker babble conditions were statistically analyzed using SPSS version

17.0 software. The mean, SD, median and the range of the SRS for each condition at different SNRs are displayed in Table 1.

Table 1: Mean, median, SD and range of the SRS in the presence of two-talker Kannada, English, mixed, interleaved babble in three SNRs two-talker mixed babble, and between two-talker Kannada babble and two-talker interleaved babble. The results of the pair-wise comparison are given in the Table 2, 3 and 4 for +3 dB SNR, 0 dB SNR and -3 dB SNR, respectively.

Condition	SNR (dB)	Mean	SD	Median	Min-Max
Two-talker Kannada babble	+3	38.00	1.414	38.00	35-40
	0	32.90	2.540	33.00	27-37
	-3	30.23	2.537	30.00	25-35
Two-talker English babble	+3	39.43	0.675	40.00	38-40
	0	37.80	0.992	38.00	36-40
	-3	34.93	2.188	35.00	28-37
Two-talker Mixed Babble	+3	38.00	.883	39.00	37-40
	0	34.60	2.479	35.00	27-39
	-3	31.68	2.777	32.00	26-35
Two-talker Interleaved Babble	+3	37.65	1.955	38.00	32-40
	0	36.60	2.098	37.00	30-40
	-3	33.25	2.771	33.50	24-37

Table 2: *Results of Wilcoxon Signed test of SRS across different babble at +3 dB SNR.*

Two talker babble condition	| Z |
English babble vs Interleaved babble	4.858*
English babble vs Kannada babble	4.562*
English babble vs Mixed babble	2.659
Interleaved babble vs Mixed babble	3.637*
Interleaved babble vs Kannada babble	0.511
Mixed babble vs Kannada	2.932*

Note: **Maximum number of key words = 40.**

The data were subjected to normality tests and the 'p' value was less than 0.05 for all conditions. Hence, non-parametric tests were used to statistically analyze the data.

3. Effect of native vs. non-native language

Friedman test was done to determine the effect of the four two-talker conditions on the SRS across these SNRs. The results of the test revealed a significant main effect of different babble condition at each SNR [+3 dB SNR ($\div 2$ (3) = 45.2, $p < 0.05$), 0 dB SNR ($\div 2$ (3) = 83.9, $p<0.05$) and -3 dB SNR ($\div 2$ (3) = 71.676, $p < 0.05$)]. At all the SNRs, except at +3 dB SNR, the rank order of sentence recognition in the presence of two-talker babble decreasing order are two-talker English babble, interleaved babble condition, mixed condition and two-talker Kannada babble condition. At +3 dB SNR, the rank order of sentence recognition in the presence of

two-talker babble in decreasing order are two-talker English Babble, two- talker mixed babble, two-talker Kannada babble and two- talker Interleaved babble.

Wilcoxon signed rank test was used for pair-wise comparison between the two-talker conditions at each SNR. The results revealed at 0 dB SNR and -3 dB SNR, there was a significant difference across all the conditions.

At +3 dB SNR, SRS in the presence of two-talker English babble was significantly higher than two-talker interleaved and Kannada babble conditions as shown in the Table 2. SRS in the presence of two-talker Kannada babble was significantly lower than two-talker mixed and English babble condition. However, SRS was significantly better in the presence of two-talker mixed condition than two-talker interleaved condition. There was no significant difference between SRS in the presence of two-talker English and mixed babble. In addition, no significant difference was found between SRS in the presence of two-talker Kannada and interleaved babble conditions.

Table 3: *Results of Wilcoxon Signed test of SRS across different babble at 0 dB SNR.*

Two talker babble condition	% Z %
English babble vs Interleaved babble	3.108*
English babble vs Mixed Babble	5.271*
English babble vs Kannada babble	5.550*
Interleaved babble vs Mixed babble	3.703*
Interleaved babble vs Kannada babble	4.701*
Mixed babble vs Kannada babble	3.380*

*$p < 0.05$

At +3 dB SNR, there was a significant difference across all conditions except between two-talker English babble and babble *$p < 0.05$

Table 4: *Results of Wilcoxon Signed test of SRS across different babbles at -3 dB SNR.*

Two talker babble condition	| Z |
English babble vs Interleaved babble	4.858*
English babble vs Mixed babble	4.562*
English babble vs Kannada babble	2.659
Interleaved babble vs Mixed babble	3.637*
Interleaved babble vs Kannada babble	0.511
Mixed babble vs Kannada babble	2.743*

*$p < 0.05$

At 0 dB SNR and -3 dB SNR, all the conditions differed significantly from one another as given in the Table 3 and 4, respectively. The SRS in the presence of two-talker English babble was significantly better than all the other babbles, whereas the SRS in the presence of two- talker Kannada babble was significantly poorer than all the other babbles. The SRS in the presence of two-talker interleaved babble was significantly poorer than two-talker mixed babble. The results of two-talker interleaved condition are elaborated below

4. Effect of Language uncertainty on SRS

The two-talker interleaved condition was used to study the effect of language uncertainty. The SRS in the presence of two-talker interleaved condition was compared to the other conditions. The Wilcoxon signed rank test results in Table 2, 3 and 5 shows that at all the SNRs, the SRS in the presence of two-talker interleaved condition was significantly poorer when compared to SRS in the presence of two-talker English and mixed condition. Further, the SRS in the presence of two-talker interleaved condition was significantly better when compared to SRS in the presence of two-talker Kannada condition at 0 dB SNR and -3 dB SNR. At +3 dB SNR, the SRS in the presence of two-talker interleaved condition was poorer than the SRS in the presence of two-talker Kannada babble; however, the difference was not significant.

5. Effect of SNR

Friedman test was done to determine the effect of SNR on SRS in each of the two-talker babble condition and the results of the test revealed a significant effect in all the four two-talker condition [two-talker English babble ($\div 2$ (2) = 76.51, $p < 0.05$), two-talker Kannada babble ($\div 2$ (2) = 69.32, $p < 0.05$), two-talker mixed babble ($\div 2$ (2) = 68.82, $p < 0.05$) and two-talker interleaved babble ($\div 2$ (2) = 57.28, $p < 0.05$)]. Further, pair-wise comparison was done using Wilcoxon signed rank test to compare the SRS across SNRs for each condition. The results of this showed that the SRS at higher SNRs yielded better recognition scores, and the SRS decreased significantly as SNR decreased in all the four two-talker conditions. Table 1 depicts the decrease in median values of SRS as SNR decreases in all the four two-talker conditions. The results of Wilcoxon signed rank test are given in the Table 5.

Table 5: *Results of Wilcoxon Signed rank test of SRS across three SNRs.*

Condition	Comparison between SNR (dB)		|Z| values
Two-talker	+3	0	5.521*
Kannada	+3	-3	5.519*
babble	0	-3	4.611*
Two-talker	+3	0	5.316*
English babble	+3	-3	5.544*
	0	-3	5.336*
Two-talker	+3	0	5.457*
mixed babble	+3	-3	5.522*
	0	-3	4.922*
Two-talker	+3	0	4.190*
interleaved	+3	-3	5.317*
babble	0	-3	5.128*

*$p < 0.05$

Discussion

The present study evaluated the performance of Kannada-English bilinguals on recognition of Kannada sentences in the presence of two-talker Kannada babble, two-talker English babble, two-talker mixed and two-talker interleaved babble. In addition to this, the effect of language uncertainty was also studied by comparing the SRS in the presence of two-talker interleaved babble and SRS in the presence of other three non-interleaved babbles.

1. Effect of native vs. non-native language

Statistical analysis of the effect of native vs. non- native babble on the sentence recognition revealed a significant better Kannada sentence recognition scores in the presence of two-talker English Babble than two-talker Kannada babble in all the SNRs. Similar results have been found in other studies (Brouwer & Bradlow, 2014; Calandruccio & Zhou, 2014; Stibbard & Lee, 2006). These studies on bilinguals have reported masking release in the presence of two-talker non-native language babble. This is attributed to the linguistic mismatch contributing to less IM in the condition where non-native babble is presented. Whereas, Anitha (2003) and Vineetha, Suma, and Nair (2013) reported no release of masking in presence of non- native language MTB (Anitha, 2003; Vineetha et al., 2013). This could be due to the fact that these studies used four to eight-talker babble, and hence, the spectral and temporal characteristic of the this would resemble a speech noise as the number of talkers used were more (Carhart et al., 1975; Cullington & Zeng, 2008; Simpson & Cooke, 2005) and hence, there was no IM.

In addition, the study done by Anitha (2003) used MTB of Indian languages of similar origin, thereby leading to lesser linguistic mismatch. The languages taken in the present study are from two different language families and are of different origin. Kannada language belongs to Dravidian family and English belongs to Indo-European family. Hence, the present study might have found masking release due to the linguistic mismatch.

However, the release of masking in the presence of non-native language babble (English babble) cannot be solely due to the difference in IM. There could be influence of energetic masking factors like F0 of the speaker, gender of the speaker and LTASS of the language, and other linguistic differences like accent (Brouwer et al., 2012; Calandruccio et al., 2010, 2013; Calandruccio & Zhou, 2014). Before attributing the linguistic mismatch as the sole factor affecting the IM, the influence of the above mentioned factors need to be ruled out. The possible effects of these factors on IM in the present study are discussed below.

a) F0 of the speakers

In the present study, the difference between the F0 of speaker of the target sentences and F0 of two speakers of Kannada babble was similar to the difference between F0 of speaker of the target sentences and both the speakers of Kannada babble. Various authors have reported a difference in the masker and target in terms of F0, and attributed the release of masking to the ability of the listener to segregate the masker and the target based on these differences rather than informational masking (Brouwer et al., 2012; Calandruccio et al., 2010, 2013; Calandruccio & Zhou, 2014).

However, in the present study, the difference in F0 between target and masker in both conditions could have aided in the segregation of the target and babble, but the difference between the scores across conditions cannot be explained based on this, because the difference between the F0 of target speaker and speakers of babble of both languages are almost same. Hence, F0 difference does not influence the difference in scores between the two-talker English and Kannada babble, rather the linguistic mismatch has played a role in the release of masking (Brungart et al., 2001; Cooke et al., 2008) when Kannada sentences were presented in the presence of two-talker English babble.

b) LTASS of the stimuli and the babble

In order to rule out EM in the present study, LTASS of different babble was analyzed. The LTASS has been reported to vary across languages (Byrne, Dillon, & Tran, 1994). This difference in LTASS can also contribute to differential EM in the speech recognition task (Calandruccio et al., 2010; Calandruccio & Zhou, 2014).Studies have reported that babble with lesser energy leads to lesser masking (Calandruccio et al., 2010, 2013).

Conversely, in the present study, only very slight differences are present in the LTASS of the four two-talker babble when compared to that the target stimuli, and hence, its contribution can be considered less influential towards the difference in the scores across two-talker babble conditions. Hence, this reveals a strong support towards the role of IM in the masking of native language in the presence of native versus non-native language babble. However, measurement of LTASS using many speakers would have been better, and normalizing the LTASS before the experimental task could have been done in order to reduce the factors contributing towards EM while studying IM (Brouwer et al., 2012).

c) Effect of accent

Calandruccio et al., (2010) reported a significant difference between mandarin accented English and native English two-talker babble at difficult SNR. In the present study, English sentences in the babble were spoken by an Indian who has been residing in the state of Karnataka. The English produced by an Indian is recognized as General Indian English (Sirsa & Redford,

2013). The Indian English is a result of accent influence of the native language. Hence, effect of the General Indian English could have more similarities acoustically and phonetically (Sirsa & Redford, 2013) with the native language i.e., Kannada language in our study.

Hence, the release of masking in the presence of two-talker English babble and more masking in presence of two-talkers Kannada babble in the present study could be attributed to difference in the lexical-semantic interference rather than acoustical and phonetic similarity contributed due to native accent influence.

2. Effect of language uncertainty on SRS

Another goal of the study was to find the effect of linguistic uncertainty of the babble on the recognition scores. The results revealed presence of both the languages in the babble (two-talker mixed and interleaved babble) condition has reduced scores than two-talker English babble and has better scores than two-talker Kannada babble. Among the mixed babble and interleaved condition, the addition of uncertainty in the interleaved condition has lead to significantly poorer scores than mixed babble condition at all SNRs.

Similar results were found in the study done by Brouwer and Bradlow (2014). They reported that the speech-in-speech recognition is sensitive to contextual variation in terms of the target-background language mismatch.

This could be due to the influence of cognitive factors like attention and working memory which affect the process of tuning in to target speech and tuning out of the speech masker. The uncertainty induces distractions which affects the selective attention. Mattys et al., (2009) reported an influence of cognitive load in terms of dual attention and divided attention cause a slowing of their cognition process and hence, affecting the speech recognition ability. This explains the reduction in the scores of sentence recognition in the presence of two-talker interleaved condition.

3. Effect of SNR

In the present study, the recognition scores were significantly higher at higher SNRs than at lower SNRs in all the babble conditions. Studies have reported that with increase in SNR, there is less interference by the masker both in terms of EM and IM, thereby leading to better scores (Brouwer et al., 2012; Brungart et al., 2001; Cooke et al., 2008; Engen & Bradlow, 2007; Van Engen, 2010; Wu et al., 2015). Another effect of SNR is that IM is seen only when both target and masker is audible to the listener. That is, at higher SNRs, the target is more audible than the masker, hence, the intensity cues facilitate in better segregation of masker and target (Engen & Bradlow, 2007) leading to less EM and IM. Whereas at difficult SNRs, the target is less audible when compared to the masker, inducing more EM and also there is less competition at the level of target speech recognition which eliminates the linguistic IM effects (Engen, 2010). Similar results have been found even in the present study.

Summary and conclusion

There has been a great influence of English on the Indian languages since the British rule and vice versa. In everyday listening situation, one is exposed to speech recognition in the presence of different language maskers and mixing of both the English and Indian languages commonly referred as code switching/code mixing. This scenario of speech-in-speech recognition gets influenced by many factors such as the language of the babble, language uncertainty of the babble and SNR.

The aim of the present study was to study the effect of native and non-native language babble on SRS of Kannada sentences, the effect of language uncertainty on the SRS of Kannada sentences, and to study the role of SNR on speech-in-speech recognition in Kannada-English bilinguals.

The SRS in 40 Kannada-English bilinguals in the presence of two-talker Kannada babble, two-talker English Babble, two-talker mixed babble and two-talker interleaved babble at +3 dB SNR, 0 dB SNR and -3 dB SNR was obtained. The influence of native language vs. non-native language was studied by comparing the SRS in the presence of two-talker Kannada babble and two-talker English Babble.

The results revealed that the linguistic mismatch did account for release in masking in the presence of two-talker English babble. In addition to this, the effect of language uncertainty was also studied by comparing the SRS in the presence of two-talker interleaved babble with the SRS in the presence of the other three non-interleaved babbles. The results revealed that addition of uncertainty deteriorated the SRS.

Further, with an increase in SNR, the scores across all four two-talker babble conditions increased systematically. However, there was no significant difference among the two-talker babble conditions seen at higher SNR (+3 dB SNR), as the target was less interfered by the two-talker babble and hence, there was lesser informational masking.

From the above results, it could be concluded that if the masker is a sentence from the native language, then the speech recognition could be poor when compared to that of non-native language. Further, unpredictable changes in the language of the masker affect the speech-in-speech recognition. This unpredictability is very common in the Indian set up and hence, this should be considered while carrying out hearing testing and hearing aid evaluation in the presence of speech babble. Further, the speech-in- speech recognition, in the present study, has been mainly due to informational

masking and less of energetic masking. In addition, the SNR is found to be a significant factor affecting speech-in-speech recognition.

Future directions for research

- The influence of one Indian language on other Indian language using two-talker babble can be studied.
- A similar study could be carried out using the same speaker for the target and babbles to control the talker differences influencing the speech-in-speech recognition.
- Various Indian languages accented English speech can be used to study the influence of Indian accent on speech recognition.

References

Anitha, R. (2003). Effect of multi-talker babble of different languages on the speech recognition scores in Kannada. Independent Project, Done at the Department of Audiology,AIISH, Mysore.

ANSI. (1991). Maximum permissible ambient noise levels for audiometric test rooms. ANSI S3.1-1991, New York: American National Standards.

Brouwer, S., & Bradlow, A. R. (2014). Contextual variability during speech-in-speech recognition. The *Journal of the Acoustical Society of America, 136*(1), EL26–EL32. doi:10.1121/1.4881322

Brouwer, S., Engen, K. J. Van, Calandruccio, L., & Bradlow, A. R. (2012). Linguistic contributions to speech-on-speech masking for native and non-native listeners/ : Language familiarity and semantic content. *The Journal of the Acoustical Society of America, 131*(2), 1449–1464. doi:10.1121/1.3675943

Brungart, D. S., & Simpson, B. D. (2004). Within-ear and across-ear interference in a dichotic cocktail party listening task: Effects of masker uncertainty. The Journal of the Acoustical Society of America, 115(1), 301–310. doi:10.1121/1.1628683

Brungart, D. S., Simpson, B. D., Ericson, M. a., & Scott,

K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. The Journal of the Acoustical Society of America, 110(5), 2527. doi:10.1121/1.1408946

Byrne, D., Dillon, H., & Tran, K. (1994). An international comparison of long-term average speech spectra. The Journal of the Acoustical Society of America, 96(4), 2108.

Calandruccio, L., Brouwer, S., Van Engen, K. J., Dhar, S., & Bradlow, A. R. (2013). Masking Release Due to Linguistic and Phonetic Dissimilarity Between the Target and Masker Speech. American Journal of Audiology, 22(1), 157–164. doi:10.1044/1059-0889(2013/12-0072)

Calandruccio, L., Dhar, S., & Bradlow, A. R. (2010). Speech-on-speech masking with variable access to the linguistic content of the masker speech. The Journal of the Acoustical Society of America, 128(2), 860–869. doi:10.1121/1.3458857

Calandruccio, L., & Zhou, H. (2014). Increase in Speech Recognition due to Linguistic Mismatch Between Target and Masker Speech: Monolingual and Simultaneous Bilingual Performance. Journal of Speech, Language, and Hearing Research, 57(3), 1089–1097. doi:10.1044/2013

Carhart, R., Johnson, C., & Goodman, J. (1975). Perceptual masking of spondees by combinations of talkers. The Journal of the Acoustical Society of America, 58(S1), S35. doi:10.1121/1.2002082

Carhart, R., Tillman, T. W., & Greetis, E. S. (1969). Perceptual Masking in Multiple Sound Backgrounds. The Journal of the Acoustical Society of America, 45(3), 694–703. doi:10.1121/1.1911445

Cooke, M., Lecumberri, M. L. G., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech. The Journal of the Acoustical Society of America, 123(1), 414–427. doi:10.1121/1.2804952

Cullington, H. E., & Zeng, F. (2008). Speech recognition with varying numbers and types of competing talkers by normal-hearing , cochlear-implant , and implant simulation subjects. The Journal of the Acoustical Society of America, 123(1), 450–461. doi:10.1121/1.2805617

Dillon, H. (1983). The effect of test difficulty on the sensitivity of speech discrimination tests. The Journal of the Acoustical Society of America, 73(1), 336. doi:10.1121/1.388815

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd, G. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. The Journal of the Acoustical Society of America, 114(1), 368. doi:10.1121/1.1577562

Engen, K. J. Van. (2012). Speech-in-speech recognition: A training study Speech-in-speech recognition: A training study. Language and Cognitive Processes, 27(7-8), 1089–1107.

doi:10.1080/01690965.2012.654644

Engen, K. J. Van, & Bradlow, A. R. (2007). Sentence recognition in native- and foreign-language multi-talker background noise. The Journal of the Acoustical Society of America, 121(1), 519–526.

Fairbanks, G. (1960). The rainbow passage. In Voice and articulation drillbook (2nd ed., pp. 124–139). New York: Harper.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. The Journal of the Acoustical Society of America, 115(5), 2246–2256.

Freyman, R. L., Helfer, K. S., & Balakrishnan, U. (2007). Variability and uncertainty in masking by competing speech. The Journal of the Acoustical Society of America, 121(2), 1040–1046. doi:10.1121/1.2427117

Geetha, C., Kumar, K. S. S., Manjula, P., & Pavan, M. (2014). Development and standardisation of the sentence

identification test in the Kannada language. Journal of Hearing Science, 4(1), 18–26.

Harini, N. V, & Shyamala, K. C. (2008). Code mixing and Code Switching in simultaneous vs successive bilingual children. Unpublished Master's Dissertation, University of Mysore.

Jain, C., Kodanath, S., Vimal, B. M., & Suresh, V. (2014). Influence of native and non-native multitalker babble on speech recognition in noise. Audiology Research, 4(1). doi:10.4081/audiores.2014.89

Jerger, J. (1970). Clinical Experience With Impedance Audiometry. Archives of Otolaryngology - Head and Neck Surgery, 92(4), 311–324.

Jones, G. L., & Litovsky, R. Y. (2008). Role of masker predictability in the cocktail party problem. The Journal of the Acoustical Society of America, 124(6), 3818–3830. doi:10.1121/1.2996336

Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. The Journal of the Acoustical Society of America, 61(5), 1337–1351. doi:10.1121/1.381436

Kidd, G., Mason, C. R., & Arbogast, T. L. (2002). Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns. The Journal of the Acoustical Society of America, 111(3), 1367. doi:10.1121/1.1448342

Mathew, M. (2012). Patterns of Code Switching in Children. Language in India, 12, 512–550.

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. Cognitive Psychology, 59(3), 203–243. doi:10.1016/j.cogpsych.2009.04.001

Neff, D. L., Dethlefs, T. M., & Jesteadt, W. (1993). Informational masking for multicomponent maskers with spectral gaps. The Journal of the Acoustical Society of America, 94(6), 3112-3126. doi:10.1121/1.407217

Savithri, S, R., & Jayaram, M. (2005). 300 words reading passages in Dravidian languages. AIISH Research Fund Project, Done at AIISH, Mysore.

Simpson, S. A., & Cooke, M. (2005). Consonant identification in N -talker babble is a nonmonotonic function of N. The Journal of the Acoustical Society of America, 118(5), 2775–2778. doi:10.1121/1.2062650

Sirsa, H., & Redford, M. a. (2013). The effects of native language on Indian English sounds and timing patterns. Journal of Phonetics, 41(6), 393–406. doi:10.1016/j.wocn.2013.07.004

Stibbard, R. M., & Lee, J.-I. (2006). Evidence against the mismatched interlanguage speech intelligibility benefit hypothesis. The Journal of the Acoustical Society of America, 120(1), 433–442. doi:10.1121/1.2203595

Van Engen, K. J. (2010). Similarity and familiarity: Second language sentence recognition in first- and second-language multi-talker babble. Speech Communication, 52(11), 943–953. doi:10.1016/j.specom.2010.05.002

Vineetha, C. V., Suma, R., & Nair, S. P. (2013). Effect of Bilingualism on Speech in Noise Perception in Young Adults. Language in India, 13(6), 799–811.

Watson, C. S., Kelly, W. J., & Wroton, H. W. (1976). Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty. The Journal of the Acoustical Society of America, 60(5), 1176–1186. doi:10.1121/1.381220

WHO. (2008). World Health Organisation. Grades of hearing impairment. Prevention of Blindness and Deafness. World Health Organization.

Wu, X., Yang, Z., Huang, Y., Chen, J., Li, L., Daneman, M., & Schneider, B. A. (2015). 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. doi:10.1044/1092-4388

Wylie, E. (2006). International Second Language Proficiency Rating Scales. Retrieved from www.balaibahasaperth.org/documents/ISLPR_scale_indonesian.pdf

Yathiraj, A., & Vijayalakshmi, C. S. (2005). Phonemically Balanced Word List in Kannada. Departmental Project, Developed in Department of Audiology, AIISH, Mysore.