



Published in final edited form as:

Speech Commun. 2010 December 30; 52(11-12): 943–953. doi:10.1016/j.specom.2010.05.002.

Similarity and familiarity: Second language sentence recognition in first- and second-language multi-talker babble^a

Kristin J. Van Engen

Department of Linguistics, Northwestern University, 2016 Sheridan Rd. Evanston, IL 60208 USA

Kristin J. Van Engen: k-van@northwestern.edu

Abstract

The intelligibility of speech in noisy environments depends not only on the functionality of listeners' peripheral auditory systems, but also on cognitive factors such as their language learning experience. Previous studies have shown, for example, that normal-hearing listeners attending to a non-native language have more difficulty identifying speech targets in noisy conditions than do native listeners. Furthermore, native listeners have more difficulty understanding speech targets in the presence of speech noise in their native language versus a foreign language. The present study addresses the role of listener language experience with both the target and noise languages by examining second-language sentence recognition in first- and second-language noise. Native English speakers and non-native English speakers whose native language is Mandarin were tested on English sentence recognition in English and Mandarin 2-talker babble. Results show that both listener groups experienced greater difficulty in English versus Mandarin babble, but that native Mandarin listeners experienced a smaller release from masking in Mandarin babble relative to English babble. These results indicate that both the similarity between the target and noise and the language experience of the listeners contribute to the amount of interference listeners experience when listening to speech in the presence of speech noise.

Keywords

speech-in-noise perception; informational masking; multi-talker babble; bilingual speech perception

1. Introduction

During everyday speech communication, listeners must cope with a variety of competing noises in order to understand their interlocutors. While it is well known that trouble understanding speech in noisy environments is a primary complaint for listeners with hearing loss, the ability to process speech in noise depends not only on the peripheral auditory system, but also on cognitive factors such as a listener's language experience (e.g.,

^aVersions of this study were presented at the Acoustical Society of America's Second Special Workshop on Cross Language Speech Perception and Variations in Linguistic Experience, Portland, OR, May 2009 and at the International Symposium on Bilingualism 7, Utrecht, The Netherlands, July 2009.

Correspondence: Kristin J. Van Engen, Department of Linguistics, Northwestern University, 2016 Sheridan Rd. Evanston, IL 60208 USA, k-van@northwestern.edu, 847-491-7020 (phone), 847-491-3770 (fax).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Mayo et al., 1997). For normal-hearing native-speaking listeners, speech intelligibility remains relatively robust even in adverse conditions. Such listeners are able to take advantage of redundancies in the speech signal (e.g., Cooke, 2006 and citations therein), as well as contextual cues at higher levels of linguistic structure, such as lexical, syntactic, semantic, prosodic, and pragmatic cues (e.g., Bradlow and Alexander, 2007). When people listen to speech in a *second* language, however, they have greater difficulty identifying speech signals (phonemes, words, sentences) in noisy conditions than do native speakers (Nábělek and Donohue, 1984; Takata and Nábělek, 1990; Mayo et al., 1997; Garcia Lecumberri and Cooke, 2006; Cooke et al., 2008; Cutler et al., 2008). Some recent data suggest that even bilinguals who acquired both languages before age 6 may have greater difficulty recognizing words in noise and/or reverberation than monolingual listeners (Rogers et al., 2006). Furthermore, when the interfering noise is also a speech signal (as in the case of multi-talker babble or a competing speaker), listeners' experience with the language of the *noise* seems to modulate their ability to process target speech: native language noise has been shown to be more detrimental than foreign-language noise for listeners' identification of native language speech targets (Rhebergen et al., 2005; Garcia Lecumberri and Cooke, 2006; Van Engen and Bradlow, 2007; Calandruccio et al., forthcoming). The present study further investigates the role of listeners' language experience in the perception of speech in noise, extending the research on both non-native speech perception and on the effects of different noise languages by examining the effects of first- and second-language noise on sentence intelligibility for listeners who are processing their second language.

Given that the effects of noise on speech perception can vary based on non-peripheral factors such as a listener's language experience, it is useful to consider the contrast drawn by hearing scientists between energetic and informational masking (see Kidd et al., 2007 and citations therein). Noise imposes energetic masking on auditory speech targets when mechanical interference occurs in the auditory periphery: components of the speech signal are rendered inaudible where there is spectral and temporal overlap between the noise and the signal. Energetic masking, therefore, is dependent on the interaction of the acoustics of the speech signal and the noise signal, and it results in the loss of acoustic and linguistic cues relevant to speech understanding. Any reduction in target speech intelligibility that is *not* accounted for by energetic masking (e.g., when both target and noise are audible, but a listener has trouble separating them) is typically described as informational masking.¹ This contrast between energetic and informational masking will be useful as we consider the effects of interfering speech on speech perception by listeners with varying language backgrounds: the relative energetic masking imposed by two noise types (English and Mandarin babble in this study) is necessarily static across listener groups, whereas the relative *informational* masking of the two noises may be modulated by listeners' language experience.

Before proceeding to a discussion of previous literature on speech perception in noise by non-native listeners, it is helpful to clarify the terms that will be used in this paper to discuss various types of noise. 'Noise' is intended to refer to any sounds in an auditory environment other than the speech to which a listener is attending. 'Masker' will be used to refer to noise that is used in experimental settings. A single interfering talker is referred to as 'competing speech', and more than one interfering talker is 'multi-talker babble' or 'babble'. 'Non-speech noise' refers to any noise that is not comprised of speech (e.g., white noise), and 'speech-shaped noise' is a type of non-speech noise that is generated by filtering broadband noise through the long-term average spectrum of speech.

¹Durlach (2006) observed that this very broad use of the term 'informational masking' reflects a lack of conceptual and scientific certainty or clarity. Kidd et al. (2007) provide a useful history and overview of the terms energetic and informational masking.

With respect to the effects of noise on non-native speech perception, several studies have shown that native listeners perform better than non-natives on speech perception tasks in stationary (i.e., without amplitude modulations) non-speech noise and multi-talker babbles containing several talkers (Mayo et al., 1997; Hazan and Simpson, 2000; Bradlow and Bent, 2002; Van Wijngaarden et al., 2002; Cutler et al., 2004; Garcia Lecumberri and Cooke, 2006; Rogers et al., 2006). These studies employed listeners from a wide range of native language backgrounds and used a variety of types of speech targets and noise. Mayo et al. used English target sentences in English 12-talker babble, and their listeners were native speakers of English, native speakers of Spanish, and early learners of both English and Spanish. Hazan and Simpson used English speech targets (VCV syllables), speech-shaped noise, and listeners who were native speakers of English, Japanese and Spanish. Bradlow and Bent (2002) used English sentences in white noise and listeners who were native speakers of English and a wide range of other languages. Cutler et al. (2004) used English CV and VC syllables as targets, English 6-talker babble, and native English and native Dutch listeners. Van Wijngaarden et al.'s (2002) targets were English, Dutch, and German sentences in speech-shaped noise, and their listeners were native speakers of Dutch, English, and German. Garcia Lecumberri and Cooke (2006) used English VCV syllables and native English and Spanish listeners. This study employed non-speech noise, English 8-talker babble, and competing speech in both English and Spanish. Only this study, therefore, investigated the effects of native and second-language noise (in the form of competing speech) on native and non-native listeners of a given language. Those results will be discussed in greater detail below.

The above studies all show, in general, poorer performance by non-native listeners on speech perception tasks in noise relative to native speakers. As noted by Cooke et al. (2008), estimates of the relative size of the native listener advantage across different levels of noise have differed across these studies. While some show that the native listener advantage increases with increasing noise levels, others show constant native listener advantages across noise levels. The size of these effects seems to be related to the nature of the speech perception task (tasks in these studies range from phoneme identification to keyword identification in sentences) and/or the precise methods used (Cutler et al., 2008). Differences aside, however, all of these studies show that non-native listeners have more difficulty identifying speech targets in noise than native listeners.

Many of the noise types used in these studies would induce primarily energetic masking (many used non-speech noises or babbles with many talkers). The specific effect of informational masking on non-native listeners of English was more recently investigated by Cooke et al. (2008). In this study, Cooke et al. explicitly investigated the roles of energetic and informational masking by comparing the effects of a primarily energetic masker (stationary non-speech noise) with a primarily informational masker (single competing talker). They found that increasing levels of noise in both masker types affected non-native listeners more adversely than native listeners. Further, a computer model of the energetic masking present in the competing talker condition showed that the intelligibility advantage for native listeners could not be attributed solely to energetic masking. The authors conclude, therefore, that non-native listeners are more affected by informational masking than are native listeners.

Cooke et al. (2008) also respond to Durlach's (2006) observation regarding the lack of specificity in the term 'informational masking' by identifying several potential elements of informational masking: misallocation of audible masker components to the target, competing attention of the masker, higher cognitive load, and interference from a "known-language" masker. In the discussion of their observed effects of informational masking on non-native listeners, then, they suggest that such listeners might suffer more from target/

masker misallocation, since their reduced knowledge of the target language (relative to native listeners) might lead to a greater number of confusions. Furthermore, they suggest that influence from the non-native listeners' native language (L1) might also result in more misallocations of speech sounds. In addition to misallocation, they also suggest that the higher cognitive load in their competing talker task (relative to the stationary noise task) may affect non-native listeners more than native listeners, given that some aspects of processing a foreign language are slower than processing a native language (Callan et al., 2004; Clahsen and Felser, 2006; Mueller, 2005). Finally, they suggest that the tracking and attention required to segregate speech signals may be compromised in non-native listeners since they have a reduced level of knowledge of the useful target language cues and/or may experience interference based on cues that are relevant for segregation of signals in the L1.

Since their study focused on a comparison of stationary noise and competing speech in the target language, Cooke et al. (2008) did not address the potential effects on non-native listeners of their final proposed aspect of informational masking: interference from a "known-language" masker. In this study, we specifically investigate this aspect of informational masking by comparing the effects of native (L1) and second-language (L2) babble on L2 sentence recognition. Native, monolingual English listeners and L2 English listeners whose L1 is Mandarin were tested on English target sentences in the presence of English 2-talker babble and Mandarin 2-talker babble. While it has been shown that English-speaking monolinguals have greater difficulty with English-language maskers as compared to foreign-language maskers (Rhebergen et al., 2005; Garcia Lecumberri and Cooke, 2006; Van Engen and Bradlow, 2007; Calandruccio et al., forthcoming), the effects of different language maskers on L2 listeners have not been thoroughly examined.

As mentioned above, the one study in which L2 listeners were tested on speech targets in L1 and L2 noise is Garcia Lecumberri and Cooke (2006). This study investigated the performance of L1 Spanish listeners on L2 (English) consonant identification in L1 and L2 competing speech, and found no difference between the listeners' performance in the two noise languages. The authors suggest that while L1 noise might be generally more difficult than L2 noise to tune out, the task of identifying L2 targets might increase interference from L2 noise, thereby eliminating the difference between the masking effects of the two languages. The present study further investigates the effects of noise language on non-native listeners by asking whether these listeners are differentially affected by L1 and L2 babble when identifying L2 sentences.

In addition to simulating an ecologically valid listening situation, sentence-length materials contain all the acoustic, phonetic, lexical, syntactic, semantic, and prosodic cues of everyday speech, and may, therefore, reveal differences between the effects of different noise types that would not be observable at the level of a phoneme identification task. With sentences, listeners are able to use redundancies in the speech signal as well as contextual linguistic cues that aid speech understanding in real-world situations. Such cues may aid perception in noise in general, but if informational masking occurs at higher levels of linguistic processing, sentence materials may also make it possible to observe differences in the effects of different noise languages. As suggested by Cutler et al. (2004), non-native listeners' difficulty in noise may reflect an accumulation of difficulties across levels of speech processing. In this case, differential effects of noise languages which may not be observable at the level of phoneme identification could be observed using materials that require the processing of more levels of linguistic structure.

By including participants who speak both babble languages (i.e., the native Mandarin group), the current study addresses another open question regarding the previously-observed differential masking by native- versus foreign-language noise on native language targets. In

the studies that showed this effect, the target language was the native language of the listeners, so the native-language babble (or competing speech) also matched the language of the target speech. It is possible, therefore, that the greater acoustic and/or linguistic similarity between the target and noise signals contributes importantly to the increased masking by native- versus foreign-language babble, regardless of the listeners' experience with the languages. With respect to acoustics, English target speech and English babble may, for example, have more similar spectral properties, leading to greater energetic masking. As for linguistic factors, English target speech and English noise share a wide range of properties (e.g., phonemes, syllable structures, prosodic features), which may make the segregation of English speech targets from English noise much more difficult—i.e., shared linguistic features may lead to greater informational masking, regardless of the native-language status of English. The present study will enable us to begin to understand, then, whether the noise language effect is primarily a *same-language* effect (i.e., similarity between target and noise leads to increased masking) or primarily a *native-language* effect (i.e., native language noise necessarily imposes more masking than another language). For the English listeners, English babble is their native language and it matches the target language. For the Mandarin listeners, however, English babble matches the target language, but Mandarin babble is their native language.

Using a different target talker from Van Engen and Bradlow (2007), we expect to replicate the finding that native English listeners have greater difficulty understanding English sentences in English versus Mandarin babble. This replication would provide additional support for the validity of the previously-observed noise language effect by showing that the effect cannot be attributed solely to the acoustic properties of a particular target voice and its interaction with the two babbles.

The performance of the Mandarin listeners in the two babble languages will allow for the comparison of interference from native-language noise versus noise in the language of the speech targets. If differential noise language effects are primarily driven by the native language status of noise, then the Mandarin babble may be more disruptive than the English babble. If such effects are primarily a result of the similarity between the target and noise languages, then the English babble may be more disruptive than the Mandarin. Finally, we may see evidence for important roles of both factors in modulating the interference that listeners experience from interfering speech. In this case, the similarity of the English babble to the target speech would make it more difficult than Mandarin babble for all listeners (due to increased energetic and/or informational masking), but the Mandarin listeners would show a smaller release from masking in Mandarin babble (i.e., a smaller performance gain in Mandarin babble relative to English babble) than the native English listeners. Crucially, this study investigates whether there are, indeed, different effects of L1 and L2 babble on L2 sentence recognition, and further, compares such effects across L1 and L2 listeners. The relative energetic masking imposed by the two noise languages is constant across the two groups, but their language experience varies and may, therefore, modulate informational masking.

Since the relative effects of English and Mandarin babble on the two listener populations is of primary interest, it was important to test both groups at signal-to-noise ratios (SNRs) that would result in similar levels of performance with respect to tolerance for energetic masking. To achieve this, listeners were tested at SNRs that were chosen relative to their individual performance on a standard speech perception test in stationary, speech-shaped noise (the Hearing in Noise Test (HINT), Nilsson et al., 1994). By normalizing the listeners according to their tolerance for energetic masking alone, the effects of two babble languages on two listener populations could be investigated.

2. Methods

The ability of each listener to understand sentences in non-speech noise (speech-shaped white noise) was measured with the Hearing in Noise Test (HINT), which employs an adaptive presentation method to estimate the SNR at which a listener can correctly repeat full sentences 50% of the time. This score was used to determine testing levels for the speech-in-babble test. Listeners were then presented with 4 blocks of 32 target sentences in 2-talker babble. Block 1 was presented at an SNR of HINT score +3 dB; Block 2 at HINT score +0 dB; Block 3 at HINT score -3 dB; and Block 4 at HINT score -6 dB. This range of SNRs was selected in order to observe performance at relatively easy and difficult noise levels and to avoid ceiling and floor effects. In each block, listeners heard a randomized set that included 16 sentences in English babble and 16 sentences in Mandarin babble (50 keywords in each). Methodological details are presented below.

2.1. Participants

2.1.1. Monolingual English listeners—Twenty-six undergraduate participants were recruited from the Northwestern University Linguistics Department subject pool and received course credit for their participation in the study. For the following reasons, 6 were omitted from the analysis presented here: 3 were bilingual; 2 had studied Mandarin; and 1 encountered a computer error. The remaining 20 participants were native speakers of English between the ages of 18 and 22 (average = 19.5), and all reported normal hearing. Four participants reported having received speech therapy in early childhood.

2.1.2. L2 English listeners—Twenty-one native speakers of Mandarin Chinese who speak English as a second language were recruited and paid for their participation. All of these participants were first-year graduate students at Northwestern University who were participating in the Northwestern University International Summer Institute, an English language and acculturation program that takes place during the month prior to the start of the academic year. One participant was excluded from analysis because she had lived in Malaysia for a number of years during childhood and therefore had had significantly different experience with English compared to the other participants, all of whom grew up in mainland China or Taiwan. The 20 included participants ranged from 22 to 32 years of age (average = 24.5), and none reported a history of problems with speech or hearing.

While English proficiency is not entirely uniform within this group, all participants had attained the required TOEFL scores for admission to the Northwestern University Graduate School and participated in the study within 3 months of their arrival in Evanston, Illinois.² In order to further characterize the L2 English participants' experience and proficiency in Mandarin and English, each person completed a lab-internal language history questionnaire and the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian et al., 2007). Table 1 provides basic information regarding the participants' English learning and proficiency.

2.2. Materials

2.2.1. Multi-talker babble—Two-talker babble was used for this experiment, largely because significant effects of babble language have been observed for sentence recognition by native English speakers using 2-talker babble. In Van Engen and Bradlow (2007), for example, significant differences were observed between English and Mandarin babble for 2-

²Most participants had not spent a significant amount of time in an English-speaking country (0–2 months), but 4 participants reported having spent 2–3 years in the U.S. at an earlier time in their lives. These listeners were included because their HINT scores fell within the range of the other participants', meaning they were not outliers with respect to the task of English sentence recognition in noise.

talker babble but not for 6-talker babble. Calandruccio et al. (forthcoming) also showed significant differences in the effects of English versus Croatian 2-talker babble. Finally, Freyman et al. (2004) showed maximal informational masking effects in 2-talker babble as compared with 3-, 4-, 6-, and 10-talker babble. While the use of a single noise type (2-talker babble only) limits investigation of the particular contributions of energetic and informational masking in this study, the primary goal is to examine the *relative* effects of two noise languages on listener groups with different experience in the two languages. The relative energetic masking imposed by the two languages is constant across both groups, so differences in their relative effects can be attributed to informational masking.

Four 2-talker babble tracks were generated in English and in Mandarin (8 tracks in total). The babble was comprised of semantically anomalous sentences (e.g. *Your tedious beacon lifted our cab.*) produced by two adult females who were native speakers of English and two adult females who were native speakers of Mandarin. The sentences were created in English by Smiljanic and Bradlow (2005) and translated into Mandarin by Van Engen and Bradlow (2007). Female voices were used for the maskers and the target in order to eliminate the variable of gender differences, which can aid listeners in segregating talkers (e.g., Brungart et al. 2001). Babble tracks were created as follows: for each talker, two sentences (a different pair for each talker) were concatenated to ensure that the noise track duration would exceed the duration of all target sentences. 100 ms of silence were added to the start of one of the two talkers' files in order to stagger the sentence start times of the talkers once they were mixed together. The two talker's files were then mixed using Cool Edit (Syntrillium Software Corporation), and the first 100 ms (in which only one talker was speaking) were removed so that the track only included portions where both people were speaking. The RMS amplitude was equalized across the finished babble tracks (4 in English; 4 in Mandarin) using Level16 (Tice and Carrell, 1998).

2.2.2. Target sentences—The target sentences come from the Revised Bamford-Kowal-Bench (BKB) Standard Sentence Test (lists 1, 5, 7–10, 15, and 21). These particular lists were selected based on their approximately equivalent intelligibility scores for normal hearing children as reported in Bamford and Wilson (1979). The BKB sentences were chosen for this study because they use a limited vocabulary that is appropriate for use with non-native listeners (see Bradlow and Bent (2002) for familiarity ratings from a highly similar population of non-native listeners). Each list contains 16 simple, meaningful English sentences and a total of 50 keywords (3–4 per sentence) for intelligibility scoring. An adult female speaker of American English produced the sentences. She was instructed to speak in a natural, conversational style, as if she were speaking to someone familiar with her voice and speech. Recording took place in a sound-attenuated booth in the Phonetics Laboratory at Northwestern University. The sentences appeared one at a time on a computer screen, and the speaker read them aloud, using a keystroke to advance from sentence to sentence. She spoke into a Shure SM81 Condenser microphone, and was recorded directly to disk using a MOTU Ultralight external audio interface. The recordings were digitized at a sampling rate of 22050 Hz with 24 bit accuracy. The sentences were then separated into individual files using Trigger Wave Convertor, an automatic audio segmentation utility developed in the Department of Linguistics at Northwestern University. The resultant files were trimmed to remove silence on the ends of the sentence recordings, and then the RMS amplitudes of all sentences were equalized using Level16 (Tice and Carrell, 1998).

2.2.3. Targets + Babble—The full set of target sentences was mixed with each of the 8 babble tracks using a utility that was developed in the Northwestern University Linguistics Department for the purpose of mixing large sets of signals. The targets and babble were mixed at a range of SNRs so that each participant could be tested at four SNRs relative to his/her HINT score (HINT +3 dB, +0 dB, -3dB, and -6dB). The various SNRs were

generated by RMS-equalizing the babble tracks at various levels relative to a static target sentence level. This basic approach to SNR manipulation has been utilized in a large number of speech-in-noise studies (e.g., Mayo et al., 1997; Sperry et al., 1997; Van Wijngaarden et al., 2002; Rogers et al., 2006) and has the advantage of maintaining a constant target level across the entire experiment. Although this method entails that the overall level of the stimuli increases as SNR decreases (that is, when the noise becomes louder with respect to the signal), previous work showed that behavioral results on this type of task were unaffected when the mixed files were re-equalized (Van Engen, 2007).

The resulting stimuli each contained a 400 ms silent leader followed by 500 ms of babble, the target and the babble, and then a 500 ms babble trailer.

2.3. Procedure

In order to determine the SNRs at which participants would be tested in the speech-in-babble experiment, the Hearing in Noise Test (HINT) was administered first (details regarding the test and its materials can be found in Nilsson et al., 1994). Using an adaptive method of presentation, the HINT estimates the SNR at which a listener can understand 50% of entire sentences in speech-shaped noise (SNR-50). (For each sentence, listeners receive an all-or-nothing score, with some allowances for errors in short, frequently reduced function words, such as ‘a’ versus ‘the’.) Listeners respond to each sentence by repeating it orally. A 20-sentence version of this test⁶ was administered diotically through Sony MDR-V700DJ earphones, and listeners were seated in a sound-attenuated booth with an experimenter. HINT thresholds were rounded to the nearest whole number for the purposes of selecting the SNRs for the speech-in-speech test.⁷

For the speech-in-babble test, listeners were seated at a desk in the sound-attenuated booth. Stimuli were presented diotically over headphones at a comfortable level. Participants were presented with a total of 132 trials—four practice sentences followed by four experimental blocks containing 32 sentences each. Each block was comprised of two BKB lists – one mixed with English babble (four of the sentences with each of the four noise tracks), the other with Mandarin babble (four of the sentences with each of the four noise tracks). Within each block, all stimuli were randomized. Listeners were instructed that they would be listening to sentences mixed with noise, and that they should write down what they heard on a provided response sheet.⁸ They were told to write as many words as they were able to understand, and to provide their best guess if they were unsure. The task was self-paced; participants pressed the spacebar on a computer keyboard to advance from sentence to sentence. They heard each sentence only once.

Before the test began, listeners were familiarized to the task and the target speaker by listening to two sentences in English babble and two sentences in Mandarin babble, all at the SNR at which they would receive the first block of testing (their HINT score +3 dB). They were told that the target talker begins speaking one-half second after the noise comes on.

⁶HINT lists 1 and 2 were used for HINT testing. Note that the original BKB sentences were used for the development of the HINT test (Nilsson et al., 1994), so there is considerable overlap between the two sets of sentences. In order to avoid the repetition of any sentence between the HINT test and the speech-in-babble test, matching sentences were removed from the selected BKB lists (1, 5, 7–10, 15, 21) and replaced with similar (matching in number of keywords and in sentence structure where possible) sentences from list 20. This amounted to a total of 7 replacements.

⁷The experiment-running software that was used for the speech-in-babble experiment required that signals and noise be mixed in advance of experimentation. Sentence targets and babble were mixed at whole-number SNRs to limit the number of required sound files to a manageable number.

⁸Written responses were used for this experiment because it can be problematic to score oral responses from non-native speakers of English. Given the difficulties native listeners have in understanding foreign-accented speech, there may be discrepancies between what a non-native participant intends to say and what the experimenter hears. Furthermore, it may be difficult to determine whether listeners are reporting words they have understood or mimicking sounds or partial words.

The experimenter played these stimuli as many times as the listener needed in order to repeat the target sentence correctly. A few listeners were unable to repeat the target after several repetitions. In these cases, the experimenter read the target to the listener, who was then given another opportunity to hear the stimulus. At this point, all listeners were able to recognize the target. After listening to the familiarization stimuli, listeners were reminded that they would be listening to the same target voice throughout the experiment.

The order of the experimental blocks was the same for every listener in that each person received the four SNRs in descending order: HINT score +3 dB, HINT score +0 dB, HINT score -3 dB, HINT score -6 dB⁹. This was done to avoid floor and ceiling effects by pitting any task or talker learning effects against SNR difficulty and possible fatigue effects¹⁰. The same two sentence lists were presented in each block for each person (e.g. lists 1 and 5 were always the target lists in Block 1), but the language of the noise mixed with each list was counterbalanced.

2.4. Data analysis

Intelligibility scores were determined by a strict keyword-correct count. Keywords with added or deleted morphemes were counted as incorrect responses, but obvious spelling errors or homophones were considered correct.

3. Results

3.1. HINT results

An unpaired, one-tailed t-test confirmed that, as predicted, monolingual English listeners had significantly lower HINT thresholds than L2 listeners ($p < .0001$) ($t = -15.0031$, $df = 27.594$, $p < .0001$). The mean scores for the two groups differed by approximately 8 dB (English mean: -2.31; Mandarin mean: 5.66). Boxplots displaying these scores are given in Figure 1.

These results replicate previous findings showing that native listeners outperform non-native listeners on speech perception tasks in energetic masking conditions (Hazan and Simpson, 2000; Bradlow and Bent, 2002; Van Wijngaarden et al., 2002; von Hapsburg et al., 2004; Garcia Lecumberri and Cooke, 2006; Rogers et al., 2006, and many others).

An investigation of the relationships between HINT scores and other measures of English experience/proficiency showed no significant correlations: age at which English acquisition began ($r = .037$, $p = .438$), years studying English ($r = -.137$, $p = .282$), and TOEFL scores ($r = -.277$, $p = .125$).

3.2. Speech-in-babble results

The mean percentage of keywords identified by the L1 listeners (monolingual English listeners) and the L2 listeners (L1 Mandarin listeners) in each noise language and at each SNR are shown in Figure 2 and given in Table 2 below.

⁹For Mandarin listeners whose HINT scores were above +7 dB ($n = 5$), speech-in-babble testing was done at SNRs of +10, +7, +4, and +1. It was determined, on the basis of other experiments run in this laboratory, that easier SNRs would make the speech-in-babble task too easy to reveal differences in performance across the two babble languages. Analysis of the performance of these individuals showed that, overall, they performed similarly to the others and did not show worse performance as a result of this limitation on the normalization scheme.

¹⁰It should be noted, too, that the babble used in this experiment is “frozen babble” (i.e., there were just four different short babble samples that listeners heard for each language during the experiment). Felty et al. (2009) compared the use of frozen and randomly varying babble (samples taken from random time points in a long babble track) on a word recognition task and found that listeners had a steeper learning curve in the frozen babble condition. This finding suggests at least one type of perceptual learning that may have occurred over the course of the experiment but would have been countered by the increasingly difficult SNRs.

Keyword identification data were assessed statistically using mixed-effects logistic regression, with subjects as a random factor and native language, babble language, SNR, and all interactions among them as fixed effects. This analysis avoids spurious results that can arise when categorical data are analyzed as proportions using ANOVAs. It also has greater power than ANOVAs and does not assume homogeneity of variances (see Jaeger (2008) for discussion of the benefits of this analysis).¹¹ Analyses were performed using R, an open-source programming language/statistical analysis environment (R development core Team 2005). The results of the regression are shown in Table 3.

The results show that the overall probability of correct keyword identification is significantly higher as SNR increases ($z = 23.06$, $p < 0.0001$) and in Mandarin versus English babble ($z = 12.05$, $p < 0.0001$). The native language background of the listener was not a significant predictor of correct response ($z = 0.27$, $p = 0.79$), showing that the method of normalizing listeners according to their HINT scores succeeded in eliminating this factor as a predictor for performance on the speech-in-babble task.

Significant interactions with language background reveal that English listeners receive a greater release from masking in Mandarin babble than do Mandarin listeners. This is supported by the significant interaction between native language and babble language ($z = -6.29$, $p < 0.0001$), which shows that Mandarin listeners generally experienced more interference from Mandarin babble than did English listeners. This interaction is particularly strong at high SNRs, as revealed by the significant interaction between these three factors ($z = -4.18$, $p < 0.0001$).

To help visualize this three-way interaction, Table 4 reports the difference in accuracy across babble languages at each SNR for each listener group. These difference scores reveal the much larger noise language effect observed in the English listeners versus the Mandarin listeners (as shown by the significant two-way interaction) and show that this effect is considerably larger at the higher SNRs (as reflected in the three-way interaction). The confidence intervals also show English listeners performed better in Mandarin versus English babble in all SNRs except the most difficult (HINT -6 dB) and that Mandarin listeners performed significantly better in Mandarin versus English noise at the easiest and most difficult SNRs (as indicated by confidence intervals that do not extend beyond 0).

In order to further investigate the effects of noise language on the non-native listeners, a mixed-effects logistic regression was also performed on the data from the Mandarin listeners only. The regression showed a significant effect of SNR ($z = 29.94$, $p < .0001$) and showed that Mandarin listeners' performance was overall better in English versus Mandarin noise ($z = 11.26$, $p < .0001$). There was also a significant interaction between the two ($z = 5.29$, $p < .0001$).

The overall analysis also revealed two-way interactions of listener group with SNR and babble language with SNR. The steeper improvement for the English versus Mandarin listeners across SNRs is reflected by a significant two-way interaction of listener group and SNR ($z = -3.50$, $p < 0.0001$). The interaction of babble language and SNR ($z = 6.68$, $p < 0.0001$) reflects the overall greater difference between noise languages at easier SNRs.

In summary, the results show that performance for both listener groups increased on the speech-in-babble task as SNR increased, and performance was generally lower in English versus Mandarin babble. Interactions involving native language background reveal that

¹¹Note that the data were also converted to percentage correct scores, transformed using the rationalized arcsine transform (Studebaker, 1985), and analyzed using a traditional repeated measures ANOVA. The results were essentially the same.

native Mandarin listeners perform relatively worse in Mandarin noise as compared with monolingual English listeners.

4. Discussion

4.1 HINT thresholds

As expected, the HINT results showed that non-native listeners require a significantly more favorable SNR (by an average difference of about 8 dB) to identify English sentences in stationary, speech-shaped noise. This finding replicates previous findings that non-native listeners have more difficulty recognizing speech in noise than do native speakers (Hazan and Simpson, 2000; Bradlow and Bent, 2002; Van Wijngaarden et al., 2002; Garcia Lecumberri and Cooke, 2006; Rogers et al., 2006 and many others).¹² Furthermore, this large and highly significant difference in scores on a standard clinical test points to the importance of taking language experience into account in the practice of audiology, and particularly in speech audiometry (von Hapsburg and Pena, 2002; von Hapsburg et al., 2004).

In addition to providing a test of our listener groups' tolerance for purely energetic masking in English sentence recognition, HINT thresholds also proved to be a useful tool for normalizing native and non-native listener performance on the speech-in-babble task. By selecting test SNRs relative to individual HINT scores, the two listener groups achieved similar performance levels on the task of English sentence recognition in 2-talker babble (as indicated by the lack of a significant effect for native language on the speech-in-speech test).

4.2 Sentence intelligibility in 2-talker babble

This study showed that, for native English speakers and L2 English speakers (L1 Mandarin), English babble was more disruptive overall to English sentence recognition than Mandarin babble. Crucially, however, it also showed that native English speakers receive a larger release from masking in Mandarin babble (a foreign language) relative to English babble than do native speakers of Mandarin. The greater overall interference from English versus Mandarin babble for both listener groups suggests that acoustic and/or linguistic similarity between the speech signal and the noise may be the most critical factor in driving noise language effects, and the greater relative interference from Mandarin babble for Mandarin-speaking listeners suggests that there is also a component of informational masking that is specifically driven by the native-language status of the noise.

For the native listeners in this study, the speech-in-speech results replicate the previously observed effect of native-/same-as-target-language versus foreign-language 2-talker babble (Van Engen and Bradlow, 2007): English babble was found to be significantly more difficult than Mandarin babble for native English listeners. The replication of this finding with a new target talker shows that the effect cannot be attributed solely to the particular acoustic or stylistic characteristics of a single target talker's voice or its interaction with the babble tracks.

In this study, the release in masking experienced by native English listeners in Mandarin versus English babble was largest at the highest tested SNR and smallest at the lowest SNR. This pattern differs from Van Engen and Bradlow (2007), which found the language effect to be largest at the most difficult SNR tested (-5 dB). In the present study, however, the difficult SNR was significantly lower than -5 dB for most listeners (as low as -10 dB for

¹²Note, of course, that the size of such differences are likely dependent (at least in part) on listeners' level of proficiency in English. The listeners in this study may be more or less proficient than listeners in other studies of non-native speech perception in noise.

the listeners with the lowest HINT thresholds). Therefore, it is likely that the higher amount of energetic masking at these lower SNRs eliminates observable linguistic informational masking effects.

While the difficulty in English versus Mandarin babble for native English listeners has primarily been considered in terms of linguistic informational masking effects, it must be noted (as pointed out by Mattys et al. 2009) that energetic masking differences between the noise languages may also exist. The English and Mandarin babble were controlled for SNR, but were not otherwise manipulated to equate, for example, long-term average speech spectrum or temporal modulation rates and depths¹³. This study avoided any signal processing that may have equated these signal-dependent factors in order to maintain the naturalness of the stimuli. This means that the two babble languages may possibly impose different amounts of energetic masking on the target sentences.

If energetic masking can account completely for the differences between English and Mandarin babble for native English listeners, then it is predicted that the effects of the two languages would be similar across listener groups. However, if the noise language effect is, indeed, driven at least in part by higher-level informational masking in the form of linguistic interference, then differential effects of noise languages on listener populations with different language experience are predicted. Furthermore, even if there are energetic masking differences across the two noise languages, differences in their relative effects on listener groups with different language experience could reveal linguistically-driven influences of informational masking. This was indeed what was observed in the present study: although their performance was lower in English babble than in Mandarin babble, native Mandarin listeners were more detrimentally affected by Mandarin babble relative to English babble than were monolingual English listeners.

With respect to non-native speech perception in noise, these results represent the first evidence that L2 babble may be more detrimental to L2 speech processing than L1 babble. As noted in the introduction, Garcia Lecumberri and Cooke (2006) did not find differential effects of L1 and L2 competing speech on L2 listeners, but several differences between these two studies may account for the different outcomes. First, the present study uses 2-talker babble, while Garcia Lecumberri and Cooke used single, competing talkers. 2-talker babble generally induces greater energetic masking, since the babble signal itself is more temporally dense than a single competing talker. It is possible that, by further reducing access to the signal, the additional energetic masking in 2-talker babble renders linguistic informational masking effects observable. It is possible that linguistic factors modulate the effects of speech noise on speech perception under relatively specific conditions. Van Engen and Bradlow (2007), for example, found no effect of babble language using a 6-talker babble.¹⁴

Noise types aside, another important distinction between these studies is that the speech perception tasks differed widely between them. This study measured L2 keyword identification in sentences, while Garcia Lecumberri and Cooke investigated L2 consonant identification. It is quite possible that the task of listening to sentence-length material is

¹³In terms of the spectral properties of these particular maskers, running t-tests did reveal differences between the languages at some frequencies; in general, however, the long-term average spectra of the English and Mandarin babbles were highly similar.

¹⁴In a detailed study of energetic and informational masking effects on speech segmentation biases, Mattys et al. (2009) also did not find differential effects of a single talker masker and an acoustically-matched modulated noise masker in a speech segmentation task (for native listeners). One of their suggestions for why language effects may emerge in 2-talker (but not in 1-talker babble) is that 2 talkers in an unintelligible language may cohere more readily for listeners, making segregation from the signal easier. This explanation may apply to the native listeners in this study, but for the non-native listeners, both maskers were intelligible. That said, their knowledge of the two languages is quite different (native and non-native), allowing, perhaps, for a tempered version of this explanation.

more susceptible to language-specific noise interference effects than is consonant identification. Sensitivity to linguistic interference from maskers may, for example, be greater when a fuller range of linguistic structures is being processed in the targets. For non-native listeners in particular, an accumulation of processing inefficiencies across levels of linguistic processing (Cutler et al., 2004) may contribute to differential sensitivity to noise languages in sentence keyword identification versus consonant identification.

Along with the issue of L2 performance in L1 and L2 babble, one of the other open questions regarding the previous finding that native English listeners are more detrimentally affected by English versus Mandarin 2-talker babble was whether this effect is primarily driven by the native-language status of English or by its greater degree of acoustic and linguistic similarity to the English targets, which may lead to greater energetic and/or informational masking. The present results from the Mandarin listeners, for whom one of the two babble maskers is native (Mandarin) and the other matches the target (English) show that, at least for the non-native listeners, interference from a 2-talker masker in the target language (English) was greater than interference from the listeners' native language (Mandarin), at least at the easiest and most difficult SNRs that were tested. This finding suggests that signal similarity (a match between target and noise languages) is at least as important as native-language status (and perhaps more) in driving noise language effects in general.

While the finding that English babble induced more interference than Mandarin babble for both listener groups points to the importance of target-masker similarity in speech-in-speech masking, the interaction with native language status also crucially implicates a significant role for language experience in informational masking: while the native and non-native groups scored similarly in English babble, particularly at the easier SNRs, the native English listeners' performance was significantly better in Mandarin babble than the non-native listeners. That is, the native Mandarin listeners had relatively more trouble 'tuning out' Mandarin babble compared to the native English listeners.

In summary, this study of speech perception in noise by native and non-native listeners has shown that *both* similarity between the target and the noise (i.e., matched language) *and* the native-language status of noise for a particular listener group contribute significantly to the masking of sentences by 2-talker babble. Future studies comparing different types of noise (e.g., competing speech, non-speech noise that is filtered or modulated to match various speech maskers) will allow for further specification of the roles of energetic and informational masking in speech perception in noise by various listener groups. In addition, experiments using other target and noise languages and other listener groups will allow for further development of our understanding of the particular role of linguistic knowledge in speech-in-speech intelligibility. For example, the typological similarity between target and noise languages may modulate the degree of interference imposed by the babble, as may the availability of semantic content of the noise to listeners. Finally, studies to investigate the level of linguistic processing at which such effects emerge (phonetic, phonological, lexical, syntactic, semantic, prosodic, etc.) will allow for a fuller understanding of the processes involved in understanding speech in the presence of speech noise.

Acknowledgments

The author thanks Ann Bradlow for helpful discussions at various stages of this project. Special thanks also to Chun Chan for software development and technical support, to Page Piccinini for assistance in data collection, and to Matt Goldrick for assistance with data analysis. This research was supported by Award No. F31DC009516 (Kristin Van Engen, PI) and Grant No. R01-DC005794 from NIH-NIDCD (Ann Bradlow, PI). The content is solely the responsibility of the author and does not necessarily represent the official views of the NIDCD or the NIH.

References

- Bamford, J.; Wilson, I. Methodological considerations and practical aspects of the BKB sentence lists. In: Bench, J.; Bamford, J., editors. *Speech-hearing tests and the spoken language of hearing-impaired children*. Academic Press; London: 1979. p. 148-187.
- Bradlow AR, Alexander JA. Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *Journal of the Acoustical Society of America* 2007;121(4):2339–2349. [PubMed: 17471746]
- Bradlow AR, Bent T. The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America* 2002;112(1):272–284. [PubMed: 12141353]
- Brungart DS, Simpson BD, Ericson MA, Scott KR. Informational and energetic masking effects in the perception of multiple simultaneous talkers. *Journal of the Acoustical Society of America* 2001;110(5):2527–2538. [PubMed: 11757942]
- Calandruccio L, Van Engen KJ, Dhar S, Bradlow AR. The effects of clear speech as a masker. *Journal of Speech, Language, and Hearing Research*. forthcoming.
- Callan DE, Jones JA, Callan AM, Akahane-Yamada R. Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory – auditory/orosensory internal models. *NeuroImage* 2004;22:1182–1194. [PubMed: 15219590]
- Clahsen H, Felser C. How native-like is non-native language processing? *Trends in Cognitive Sciences* 2006;10:564–570. [PubMed: 17071131]
- Cooke M. A glimpsing model of speech perception in noise. *Journal of the Acoustical Society of America* 2006;119(3):1562–1573. [PubMed: 16583901]
- Cooke M, Garcia Lecumberri ML, Barker J. The foreign language cocktail party problem: energetic and informational masking effects on non-native speech perception. *Journal of the Acoustical Society of America* 2008;123(1):414–427. [PubMed: 18177170]
- Cutler A, Garcia Lecumberri ML, Cooke M. Consonant identification in noise by native and non-native listeners: effects of local context. *Journal of the Acoustical Society of America* 2008;124(2):1264–1268. [PubMed: 18681612]
- Cutler A, Webber A, Smits R, Cooper N. Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America* 2004;116(6):3668–3678. [PubMed: 15658717]
- Durlach N. Auditory masking: Need for improved conceptual structure. *Journal of the Acoustical Society of America* 2006;120(4):1787–1790. [PubMed: 17069274]
- ETS. TOEFL Internet-based Test: Score Comparison Tables. Educational Testing Services; 2005.
- Felty RA, Buchwald A, Pisoni DB. Adaptation to frozen babble in spoken word recognition. *Journal of the Acoustical Society of America* 2009;125(3):EL93–EL97. [PubMed: 19275281]
- Freyman RL, Balakrishnan U, Helfer KS. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *Journal of the Acoustical Society of America* 2004;115(5):2246–2256. [PubMed: 15139635]
- Garcia Lecumberri ML, Cooke M. Effect of masker type on native and non-native consonant perception in noise. *Journal of the Acoustical Society of America* 2006;119(4):2445–2454. [PubMed: 16642857]
- Hazan V, Simpson A. The effect of cue-enhancement on consonant intelligibility in noise: speaker and listener effects. *Language and Speech* 2000;43(3):273–294. [PubMed: 11216296]
- Jaeger TF. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 2008;59:434–446. [PubMed: 19884961]
- Kidd, G.; Mason, CR.; Richards, VM.; Gallun, FJ.; Durlach, NI. Informational Masking. In: Yost, WA.; Popper, AN.; Fay, RR., editors. *Auditory Perception of Sound Sources*. Springer; US: 2007. p. 143-189.
- Marian V, Blumenfeld HK, Kaushanskaya M. The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *J Speech Lang Hear Res* 2007;50(4):940–967. [PubMed: 17675598]

- Mattys SL, Brooks J, Cooke M. Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*. 2009
- Mayo LH, Florentine M, Buus S. Age of Second-Language Acquisition and Perception of Speech in Noise. *Journal of Speech, Language and Hearing Research* 1997;40:686–693.
- Mueller JL. Electrophysiological correlates of second language processing. *Second Language Research* 2005;21(2):152–174.
- Nábělek AK, Donohue AM. Perception of consonants in reverberation by native and non-native listeners. *Journal of the Acoustical Society of America* 1984;75:632–634. [PubMed: 6699302]
- Nilsson M, Soli SD, Sullivan JA. Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America* 1994;95(2):1085–1099. [PubMed: 8132902]
- R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2005.
- Rhebergen KS, Versfeld NJ, Dreschler WA. Release from informational masking by time reversal of native and non-native interfering speech (L). *Journal of the Acoustical Society of America* 2005;118(3):1–4.
- Rogers CL, Lister JJ, Febo DM, Besing JM, Abrams HB. Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics* 2006;27:465–485.
- Smiljanic R, Bradlow AR. Production and perception of clear speech in Croatian and English. *Journal of the Acoustical Society of America* 2005;118(3):1677–1688. [PubMed: 16240826]
- Sperry JL, Wiley TL, Chial MR. Word recognition performance in various background competitors. *Journal of the American Academy of Audiology* 1997;8(2):71–80. [PubMed: 9101453]
- Studebaker GA. A “rationalized” arcsine transform. *Journal of Speech and Hearing Research* 1985;28:455–462. [PubMed: 4046587]
- Takata Y, Nábělek AK. English consonant recognition in noise and in reverberation by Japanese and American listeners. *Journal of the Acoustical Society of America* 1990;88(2):663–666. [PubMed: 2212289]
- Tice, R.; Carrell, T. Level 16. Lincoln, Nebraska: University of Nebraska; 1998.
- Van Engen KJ. A methodological note on signal-to-noise ratios in speech research. *Journal of the Acoustical Society of America* 2007;122(5):2994.
- Van Engen KJ, Bradlow AR. Sentence recognition in native- and foreign-language multi-talker background noise. *Journal of the Acoustical Society of America* 2007;121(1):519–526. [PubMed: 17297805]
- Van Wijngaarden S, Steeneken H, Houtgast T. Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America* 2002;111:1906–1916. [PubMed: 12002873]
- von Hapsburg D, Champlin CA, Shetty SR. Reception thresholds for sentences in bilingual (Spanish/English) and monolingual (English) listeners. *Journal of the American Academy of Audiology* 2004;14:559–569.
- von Hapsburg D, Pena ED. Understanding bilingualism and its impact on speech audiometry. *Journal of Speech, Language, and Hearing Research* 2002;45:202–213.

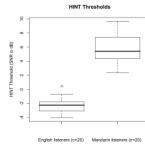


FIGURE 1.

HINT threshold scores (the SNR at which a listener can identify whole sentences 50 percent of trials) for the native English listeners and the native Mandarin listeners. The center line on each boxplot denotes the median score, the edges of the box denote the 25th and 75th percentiles, and the whiskers extend to data points that lie within 1.5 times the interquartile range. Points outside this range appear as outliers.

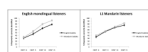


FIGURE 2. Mean intelligibility scores expressed as percentage of correct keyword identifications for native English listeners (left) and native Mandarin (L2 English) listeners (right). Error bars represent standard error.

TABLE 1

Native Mandarin participants: English learning and proficiency information.

Native Mandarin listeners	Mean (SD)
Age of initial English learning (age in years)	11.3 (1.3)
Years of formal English training	9.6 (2.2)
Self-reported TOEFL score (iBT) ³	106.2 (6.4)
Self-reported proficiency – understanding English	6.1 (1.7) ⁴
Self-reported proficiency – speaking English	5.8 (1.7) ⁵

³1 participant did not report any TOEFL scores; 3 reported paper-based test scores and 1 reported a computer-based score. All scores were converted to internet-based test scores using comparison tables from the test administration company (ETS, 2005).

⁴As rated by participants on a scale from 0 (*none*) to 10 (*perfect*). A rating of 6 indicates *slightly more than adequate*.

⁵As rated by participants on a scale from 0 (*none*) to 10 (*perfect*). A rating of 6 indicates *slightly more than adequate*.

TABLE 2

Means and standard errors of English and Mandarin listeners' recognition scores in each noise condition (% keywords identified).

Listeners	Eng+3	Man+3	Eng+0	Man+0	Eng-3	Man-3	Eng-6	Man-6
English	Mean (Std. Error)	78.2 (3.17)	91.8 (1.62)	64.5 (3.85)	78.9 (2.71)	44.6 (3.95)	53.5 (3.75)	30.4 (3.89)
Mandarin	Mean (Std. Error)	75.3 (2.98)	82.4 (1.89)	67 (2.08)	69.8 (2.62)	53.3 (3.00)	54.9 (2.31)	34.3 (2.97)
								39.8 (2.51)

TABLE 3

Summary of logistic regression on probability of correct response including participant as random intercept (Overall intercept: 0.606; St.Dev. of participant intercepts: 0.529).

Predictor	Estimate (B)	Standard Error (SEB)	Odds ratio (e ^B)
SNR	0.259***	0.011	1.295
Babble Language (Mandarin vs. English)	0.784***	0.065	2.190
Native Language (Mandarin vs. English)	0.048	0.177	1.049
Native Language * Babble Language	-0.544***	0.087	0.580
Native Language * SNR	-0.054***	0.015	0.947
Babble Language * SNR	0.115***	0.017	1.121
Native Language * Babble Language * SNR	-0.096***	0.023	0.908

Significance values:

* p < .05;

** p < .001;

*** p < .0001

TABLE 4

Mean differences with 95% confidence intervals for keywords identified in Mandarin - English babble (in percentage correct).

% keywords identified in Mandarin - English babble	HINT + 3dB	HINT + 0dB	HINT -3dB	HINT -6dB
English listeners	13.6 (7.3–20.9)	14.4 (9.1–19.7)	8.9 (2.8–15.0)	2.7 (–2.4–7.8)
Mandarin listeners	7.1 (1.4–12.8)	2.8 (–.5–6.1)	1.6 (–2.9–6.0)	5.5 (1.0–10.0)