Published in final edited form as:

J Phon. 2013 September ; 41(5): . doi:10.1016/j.wocn.2013.06.003.

Listening with a foreign-accent: The interlanguage speech intelligibility benefit in Mandarin speakers of English

Xin Xie^{a,b,*} and Carol A. Fowler^{a,b}

^aUniversity of Connecticut, Storrs, CT 06269, USA

^bHaskins Laboratories, New Haven, CT 06511, USA

Abstract

This study examined the intelligibility of native and Mandarin-accented English speech for native English and native Mandarin listeners. In the latter group, it also examined the role of the language environment and English proficiency. Three groups of listeners were tested: native English listeners (NE), Mandarin-speaking Chinese listeners in the US (M-US) and Mandarin listeners in Beijing, China (M-BJ). As a group, M-US and M-BJ listeners were matched on English proficiency and age of acquisition. A nonword transcription task was used. Identification accuracy for word-final stops in the nonwords established two independent interlanguage intelligibility effects. An interlanguage speech intelligibility benefit for listeners (ISIB-L) was manifest by both groups of Mandarin listeners outperforming native English listeners in identification of Mandarin-accented speech. In the benefit for talkers (ISIB-T), only M-BJ listeners were more accurate identifying Mandarin-accented speech than native English speech. Thus, both Mandarin groups demonstrated an ISIB-L while only the M-BJ group overall demonstrated an ISIB-T. The English proficiency of listeners was found to modulate the magnitude of the ISIB-T in both groups. Regression analyses also suggested that the listener groups differ in their use of acoustic information to identify voicing in stop consonants.

1. Introduction

Non-native speech has been widely reported to be harder for native speakers to understand than native speech (e.g., Bent & Bradlow, 2003; Flege, 1989). Considerable evidence has shown that speakers' second language (L2) proficiency affects intelligibility (Flege, MacKay & Meador, 1999; Munro & Derwing, 1995). However, foreign-accented speech is not inferior in intelligibility to all listeners.

In particular, the language background shared between the talker and listener and individual language experience affect the intelligibility of foreign-accented speech. Bent and Bradlow (2003) coined the term "interlanguage speech intelligibility benefit" (ISIB) to represent a situation in which L2 learners identify foreign-accented speech with equal or greater accuracy than they identify native speech. Their use of the term "interlanguage" refers to the possibility that the benefit reflected a language shared between L2 speakers and listeners with properties both of L1 and L2. In Bent and Bradlow's study, the number of key words identified in a sentences-in-noise perceptual task was used as the measure of intelligibility. The researchers reported that both Chinese and Korean listeners showed an ISIB for talkers with whom they shared the first language (L1). Similar findings are reported by other

investigators (e.g., Smith & Rafiqzad, 1979; Van Engen et al., 2010; see Leikin, Ibrahim, Eviatar, & Sapir, 2009 for Hebrew as L2). In particular, Bent, Bradlow, and Smith (2008) examined temporal patterns involved in the voicing of final consonants in English. They found that non-native talkers may adopt different strategies (for example, lengthening the vowel before voiced tokens or not) in both production and perception when compared to native talkers. Furthermore, these strategy differences might account, at least partially, for the differential performance of native vs. non-native listeners when they are listening to foreign-accented speech.

Although Bent and Bradlow (2003) ascribed the benefit to an interlanguage, it is possible that the ISIB does not occur exclusively in cases in which L1 is shared by the talker and listener. For instance, experience with a particular foreign-accent improves its intelligibility (Bradlow & Bent, 2008; Hanulíková & Weber, 2012; Jongman, Wade, & Sereno 2003), and L2 speakers of a given L1 are typically more experienced hearing their own foreign accent in L2 than speakers of other languages.

Whereas Bent and Bradlow had counted as a “benefit” identification of L2 speech that was either equal to or superior to identification of native speech, Stibbard and Lee (2006) later proposed that “benefit” should refer only to situations in which L2 listeners or speakers outperform native listeners or speakers. We adopt this usage in the current study.

1.1. Goal 1: To confirm that the ISIB-L and ISIB-T are independent phenomena

Hayes-Harb, Smith, Bent, and Bradlow (2008) distinguished between two types of ISIB. As we show in Fig. 1, ISIB can refer to: (1) a benefit for non-native listeners over native listeners when speech is non-native (thus, an ISIB for listeners, or an ISIB-L; e.g., Hayes-Harb et al., 2008; Major, Fitzmaurice, Bunta, & Balasubramanian 2002); or (2) a benefit in intelligibility for non-native listeners when speakers are non-native versus native (thus, an ISIB for talkers, or an ISIB-T; e.g., Bent & Bradlow, 2003; Bent et al., 2008; Stibbard & Lee, 2006). In a first attempt to distinguish these two types of benefit, Hayes-Harb et al. (2008) found no evidence for an ISIB-T. The two types of benefit have not been systematically recognized as independent in the literature. One goal of the current study is to provide further evidence in this regard.

1.2. Goal 2: To examine the role of L2 proficiency in modulating the ISIB

Some studies have raised doubts about the generality of ISIBs across languages. For each type of ISIB, there is conflicting empirical evidence. With respect to the ISIB-L, several studies have shown that non-native English listeners, for example, native Spanish (Imai, Walley, & Flege 2005; Major et al., 2002), native Japanese (Munro, Derwing, & Morton 2006) and native Mandarin (Hayes-Harb et al., 2008) listeners recognized foreign-accented English speech better than native English listeners when the speech was produced by a speaker who shared their L1. In contrast, Munro et al. (2006) failed to find a benefit for Cantonese listeners, and Major et al. (2002) even demonstrated a disadvantage for Mandarin listeners. With respect to the ISIB-T, evidence is scant: van Wijngaarden et al. (2002) showed that Dutch listeners, who spoke both German and English as L2s, exhibited an intelligibility benefit for German speech produced by a native Dutch speaker compared to a native German speaker but not for Dutch-accented English speech compared to native English speech.

A possible explanation for the inconsistency in these past studies is that the L2 proficiency of listeners and speakers may modulate intelligibility benefits. For example, an ISIB-L was found only with low proficient Mandarin listeners listening to low proficient speech (Hayes-Harb et al., 2008). In the study of van Wijngaarden et al. (2002), Dutch listeners, who

showed an ISIB-T with Dutch-accented German compared to native German but not with Dutch-accented English compared to native English, were highly-proficient in English but had low proficiency in German. These results perhaps suggest that an ISIB-T is found only among low-proficient listeners. However, other studies have found no evidence for an ISIB-T (Hayes-Harb et al., 2008; Stibbard & Lee, 2006) regardless of participants' L2 proficiency. Thus, there is limited evidence available on the ISIB, and existing empirical studies of L2 proficiency as a factor have not provided a consistent picture of its role in modulating the ISIB. Therefore, a second goal of the current study is to examine the role of L2 proficiency in modulating the ISIB.

1.3. Goal 3: To assess the effect of the language environment on the ISIB

Another factor that may affect the ISIB is the language environment. A few months of exposure to a new speaking environment can cause a shift in non-native speech production (Sancier & Fowler, 1997). In natural situations, evidence suggests that long-term perceptual adaptations can occur after six months in an L2-speaking country, even for late learners (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada 2004; MacKain, Best, & Strange 1981). Length of residence, age of acquisition, relative quantity and quality of input from native L2 speakers, and ratio of L1/L2 usage have all been identified as contributing to L2 immersion learning (see Flege, 2007 for a review). However, because most studies have been done with L2 learners in an English-speaking country, these factors are usually highly correlated.

Access to native English speech can be quite limited for L2 speakers in their native environment (Bolton, 2002; Zhang & Elder, 2011). Typical L2 learners of English in China, the population of interest in our study, are exposed only to Mandarin- or other Chinese language-accented English during their years of classroom learning, even in metropolitan cities such as Beijing. This may shape their L2 phonological representations differently from those exposed to native English, even though classroom learners can achieve high proficiency in English.

Pinet, Iverson, and Huckvale (2011) compared perception of English sentences in noise produced by native and non-native speakers among French listeners who were tested in the UK and France separately. Inexperienced listeners in France were more accurate with French-accented English than with native English but experienced listeners in the UK were more accurate with native English speech than accented speech. The third goal of the current study is to further control L2 proficiency among listeners and directly examine the language environment as a factor in the ISIB. By recruiting two groups of Mandarin-speaking listeners who are matched on age, L2 learning method, and age of acquisition but who live in different language environments, one English speaking and one Mandarin speaking, we can assess the effect of the language environment on L2 speech perception and its contribution to the ISIB. We can also ask whether the two groups of Mandarin speakers differ in whether they manifest an ISIB-L and/or an ISIB-T.

We can further look at the language use of the two Mandarin L1 groups to explain any differences in the ISIB between the groups. We will use a language questionnaire to examine language background, learning and usage patterns of L1 and L2, and other learner-related variables. By examining these variables in two different groups, we aim to explore what about the language environment affects the ISIB.

1.4. Goal 4: To link acoustic characteristics of native speech and foreign-accented speech with the ISIB

A final goal of our study is to understand whether the native and non-native speakers who provide the recordings for our listeners differ in the acoustic specification of the stop-voicing contrasts, whether the three groups of listeners attend differently to the different kinds of acoustic information for stop voicing that the speakers may provide, and whether these differences give rise to the observed ISIB effects.

Studies show that different acoustic information is used in systematically different manners by native and non-native listeners to differentiate contrasting phonetic segments (e.g., Flege, 1989; Iverson et al., 2003). For example, Flege (1989) found that removing final release bursts had a substantial negative effect on Mandarin listeners in discriminating the word-final English /t/-/d/ contrast, even when preceding vowel information was present; in contrast, the manipulation had no effect on native-English listeners. Iverson et al. (2003) looked at acoustic cues used by Japanese, German and American English adults to categorize /la/-/ra/ token in English. Japanese listeners were observed to be most sensitive to F2, an acoustic cue that is irrelevant to the English /l/-/r/ categorization while German and English listeners showed more sensitivity to F3, a critical cue that distinguishes /l/ from /r/. These differences may reflect listening strategies developed for the L1 and may explain the difficulty encountered by L2 learners, for example, Mandarin listeners distinguishing English word-final stops or Japanese listeners acquiring /l/-/r/ contrast.

In Mandarin, stops are all voiceless and differ in aspiration and place of articulation. Moreover, they occur only in word-initial position. Thus, unlike in English, voicing is not an important feature for distinguishing stop consonants in Mandarin. These differences from English are likely to give rise to strategies for identifying stop categories by aspiration that differ from strategies used by English listeners to identify stop categories by their voicing categories. Several studies have provided evidence that Mandarin-speaking listeners tend to pay attention to information in the final consonant rather than the vowel lengthening cue to differentiate word-final consonants in English (Bent et al., 2008; Flege, 1989).

In our study, we examine whether our English and Mandarin speakers signal stop voicing in different ways, whether members of our different listener groups attend differently to the acoustic specifications of the two speakers, and whether these differences can explain the observed ISIBs. Relations between stop consonant intelligibility scores and the temporal-acoustic characteristics of stimulus items will be assessed by correlation and regression analyses.

2. Methods

2.1. Participants

Three groups of listeners participated in the study: native-English speakers (NE), native-Mandarin speakers in the US who speak English as an L2 (M-US); native-Mandarin speakers in Beijing, China who also speak English as an L2 (M-BJ). NE participants were undergraduates at the University of Connecticut (UConn). Of 51 NE participants tested, only data from 34 monolingual English speakers who had no experience with any foreign languages were analyzed. An initial group of 28 Mandarin-speaking participants was tested, all undergraduates or graduate students at UConn. The data from 3 M-US listeners who were early L2 learners (who arrived in the US before the age of 8 years) were excluded, leaving data from 25 M-US participants. For purposes of comparing the two groups of L2 English speakers, this is conservative, because it will reduce differences between the groups. The other group of Mandarin speakers included undergraduates or graduate students at Beijing Normal University, China (M-BJ: $n = 30$). All participants received academic credits or

monetary rewards for participation. No Mandarin speakers reported speaking any languages other than Mandarin Chinese and English. (Some spoke a different dialect of Chinese from birth, but they used mostly Mandarin when they spoke Chinese). None reported a hearing disorder.

A self-report survey was used to collect information about participants' language background: age of acquisition (AoA) for L1 and L2 (if applicable), length of residence in the US (months, if any); the frequency of L1 and L2 (for the M-US and M-BJ groups only) use on a daily basis (hours per day), separately for native English listening and Mandarin-accented English listening. A sample item was: "During the course of a typical day, I hear: _ hrs of native-accented English; _ hrs of Mandarin-accented English." We also asked participants to report the type of setting (e.g., with friends, family members, playing sports, doing schoolwork, etc.) in which they tended to use these languages. Standardized English test scores (TOEFL) were also collected from the Mandarin participants who were willing to provide them. 23 out of 30 M-BJ listeners and 21 out of 25 M-US listeners reported scores on TOEFL or equivalent standardized tests. TOEFL scores were collected to provide an objective measure of Mandarin listeners' overall English proficiency.

Previous studies (e.g., Pinet & Iverson, 2010; Pinet et al., 2011) have compared the performance of L2 English speakers in an English-speaking country (for example, UK), with relatively high L2 proficiency, with L2 speakers in other countries (for example, France), usually with lower overall L2 proficiency. In the current study, TOEFL exam scores indicated that overall all the Mandarin listeners exhibited good English language skills. There was no significant statistical difference between the TOEFL scores of the two groups. Thus, the overall English proficiency of the Mandarin listeners was controlled, to examine the effect of the language environment on English perception.

For the M-US group, there was no significant correlation between length of residence in the US ($M = 19.8$ months, $SD = 18$, ranging from 3 to 66 months) and measures of speech perception accuracy, whether the speech was native-English or Mandarin-accented. All other measures obtained from the questionnaire were comparable between the M-US and M-BJ groups (for whom length of residence was uniformly 0 months) (see Table 1). T-test results revealed no difference between the two groups on any of the measures at the .05 significance level.

2.2. Stimuli

Two lists of 138 "English" monosyllabic nonwords with stop consonants in word-final position were constructed (e.g., *ved*, *zib*, *sheeg*, *sut*, *doop*, *roak*) for the nonword transcription task. In each list there were 20–26 tokens for each final consonant, resulting in 70 items ending in voiced stops (/b/, /d/, /g/) and 68 ending in voiceless stops (/p/, /t/, /k/). Only one instance of each nonword was presented. The rationale for using isolated nonwords rather than sentences as in the study of Bent and Bradlow (2003) was twofold: first, to diminish lexical effects that might give the NE listeners an advantage; second, to make the task cognitively easier than a sentence task, so that any differences observed between the groups can be more confidently attributed to perceptual factors. Two age-matched female speakers read the nonwords from a list and were recorded in a sound-proof booth. One was a native-English speaker (NE); the other was a native-Mandarin speaker (M-US) who was highly-proficient in English and who came to the US from China 15 months before she was recorded. She had been working as a teaching assistant since her arrival. (To serve as a teaching assistant at UConn, a non-native speaker of English has to demonstrate good intelligibility.) This speaker began learning English 15 years before the time of testing. We used such a highly proficient speaker because previous studies found an ISIB-T only with highly-proficient speakers (e.g., Bent & Bradlow, 2003). The recorded stimuli were

randomized and put into two lists. Both lists included nonwords produced by both speakers. However, items were counterbalanced by speaker such that they were presented in the same order in the lists but were produced by different speakers. For example, if the nonword *ved* in list 1 was produced by the NE speaker, it was produced by the M-US speaker in list 2. Stop consonant ([b]/[p], [g]/[k], [d]/[t]) identification accuracy in word-final position was computed as a percentage, with listeners' spelling responses coded as correct if they matched the consonant intended by the speaker.

2.3. Procedure

Participants were seated in a sound-attenuated room in front of a computer screen. They were told that they would be participating in a speech study in which they would hear recordings of some nonwords played through headphones. Their task was to type the nonword that they heard. The loudness of stimuli was adjusted by each participant to a comfortable listening level. On each trial, participants pressed a button to play the target nonword; they entered their response using a computer keyboard. Participants could replay each item as many times as they wanted. Four practice trials were given to ensure participants' understanding of the procedure and to ensure that they could clearly hear the stimuli. Participants were randomly assigned to receive one of the two counterbalanced lists so that each participant heard each target word only once, by one speaker or the other. We emphasized that they should type what they heard and not try to make items into a familiar word. After the nonword transcription task, participants completed the survey of their language background.

3. Results

3.1. Goal 1: To confirm that the ISIB-L and ISIB-T are independent phenomena

Listeners' accuracy in identifying the final stops is shown in Fig. 2. The left side shows the identification results for voiced coda consonants produced by each speaker (dark solid bars for the NE speaker and light striped bars for the M-US speaker), averaged over listeners within each group. The right side shows the accuracy on voiceless tokens.

We compared the groups' accuracies in a $3 \times 2 \times 2$ ANOVA design with listener group (NE vs. M-US vs. M-BJ) as a between-group factor, and speaker (NE vs. M-US) and voicing (voiced vs. voiceless coda stop consonant) as within-group factors. The dependent measure is the percentage of correct identifications. The ANOVA revealed a significant main effect of listener group ($F(2,86) = 3.134$, $p < .05$, partial $\eta^2 = .068$) with English listeners being overall more accurate than Mandarin listeners. There was also a significant main effect of speaker ($F(1,86) = 382.443$, $p < .001$, partial $\eta^2 = .816$) with performance numerically higher for the native speaker. Finally, there was a significant main effect of voicing ($F(1,86) = 326.073$, $p < .001$, partial $\eta^2 = .791$) with performance near ceiling on voiceless consonants, but well below ceiling on voiced consonants. Interpretation of these main effects was modulated by several significant interactions: between speaker and listener group ($F(2,86) = 16.773$, $p < .001$, partial $\eta^2 = .281$), voicing and speaker ($F(1,86) = 157.542$, $p < .001$, partial $\eta^2 = .647$) and a three-way interaction ($F(2,86) = 32.789$, $p < .001$, partial $\eta^2 = .433$). The three-way interaction was significant because English listeners showed a substantial accuracy difference depending on the speaker, whereas the two Mandarin groups showed smaller differences; moreover, this pattern held only for the voiced stops, because performance on voiceless consonants was at ceiling.

The speaker $\times$ listener group interaction is directly relevant to our question whether our listeners showed ISIB-Ls and ISIB-Ts. We followed up our overall analyses with planned comparisons. To test for an ISIB-L, we performed planned multiple comparisons using

Bonferroni tests comparing performance (that is, overall accuracy on the nonword task, including both voiceless and voiced tokens) among listener groups for each speaker. These tests revealed that, for the M-US speaker, both M-BJ listeners ($M = .78$, $SD = .11$) and M-US listeners ($M = .76$, $SD = .10$) performed significantly better than the NE listeners ($M = .69$, $SD = .07$), both $ps < .01$. This provides clear evidence for an ISIB-L for both Mandarin listener groups; for the NE speaker, the reverse pattern was observed: the NE listeners ($M = .94$, $SD = .04$) performed significantly better than either of the other two listener groups (M-US: $M = .79$, $SD = .13$; M-BJ: $M = .75$, $SD = .10$, both $ps < .001$) on the native English speech.

Importantly, although M-BJ listeners performed numerically better than the M-US listeners with M-US speech, and the M-US listeners performed better than the M-BJ listeners with NE speech, these differences were not significant. This suggests that, as a group, the M-BJ listeners were matched with M-US listeners on accuracy level.

Paired comparisons were also conducted to look for an ISIB-T. For the M-US listener group, there was no significant difference in accuracy between the two speakers ($t(24) = -2.029$, $p = .054$); for the M-BJ group, the M-US speaker was significantly more intelligible than the NE speaker ($t(29) = 2.28$, $p < .05$). Thus, an ISIB-T was observed for the M-BJ group but not the M-US group. In contrast, for the NE group, the NE speaker was significantly more intelligible than the M-US speaker ($t(33) = -25.122$, $p < .001$).

The interaction between voicing and listener group was overall nonsignificant. For all three listener groups, the voiceless stops were significantly easier to identify than their voiced counterparts ($p < .001$). Considering voiced stop consonants alone, with NE speech, NE listeners were significantly more accurate than M-US and M-BJ listeners ($F(2, 86) = 33.74$, $p < .001$); with M-US speech, a reverse pattern was found ($F(2, 86) = 9.394$, $p < .001$); between the M-US and M-BJ groups, the difference was nonsignificant for either speaker.

3.2. Goals 2 and 3: The roles of language proficiency and language environment

Demographic information collected from the language questionnaire indicated comparable overall English proficiency for two Mandarin listener groups. Therefore, we focused on participants' individual abilities to gauge whether their English proficiency is related to the fact that only the M-BJ group showed an ISIB-T.

3.2.1. Goal 2: To examine the role of L2 proficiency in modulating the ISIB—

We conducted the first set of analyses on learner-related variables among Mandarin-native listeners to address the second goal of this study to examine the role of L2 proficiency in modulating the ISIB, here, the ISIB-T. In the present study, unlike previous studies (e.g., Hayes-Harb et al., 2008), we did not use an accentedness judgment task to gauge the phonological proficiency of our listeners. Instead, we used listeners' perceptual proficiency in English as our central measure of proficiency. Mandarin-speaking listeners from both groups were included in the analysis. We classified listeners as highly-proficient (HP) if their overall accuracy on NE speech was higher than .85. (The average accuracy of the NE listener group was $.94 \pm .04$.) Medium-proficient listeners (MP) had accuracies between .75 and .85. Low-proficient listeners (LP) had accuracies below .75. With this operational definition, the proficiency of listeners is naturally related to the magnitude of the ISIB-L, so the following analyses were on the ISIB-T.

As our measure of the magnitude of the ISIB-T, we calculated a difference score by subtracting accuracy on NE speech from accuracy on M-US speech. Accordingly, an ISIB-T effect was indexed by a positive accuracy difference. A t-test indicated that the scores differed for the two Mandarin listener groups: M-US ($-.03$) and M-BJ ($.04$) group ($t(53) =$

3.031, $p < .01$). This confirmed the earlier analysis, which showed an ISIB-T only for the M-BJ group.

We next examined whether the outcome was modulated by individual proficiency. The percentages of the listeners falling into each proficiency category were 44%, 36% and 20% for the M-US group and 50%, 33.3%, and 16.7% for the M-BJ group, from low to high. The frequencies of participants in each proficiency category did not differ between the groups, $\chi^2(2) = .215$, $p = .898$, implying that the groups were matched on proficiency.

A 2 (listener group: M-US and M-BJ) $\times$ 3 (proficiency level: HP, MP and LP) factorial analysis with the ISIB-T difference score as the dependent measure revealed a gradient effect of proficiency within each demographic group, M-US and M-BJ. That is, there was a significant effect of proficiency ($F(2, 49) = 3.273$, $p < .05$, partial $\eta^2 = .118$) such that the lower the proficiency, the larger the ISIB-T effect (LP group (.04) > MP group (.01) > HP (−.04)). Post-hoc tests revealed a significant difference only between the HP and LP groups ($p < .05$), indicating gradience. This result is consistent with that of van Wijngaarden et al. (2002) in which the ISIB-T was observed among low-proficient Dutch listeners of German. At each proficiency level, the average ISIB-T effect was numerically larger in the M-BJ than M-US group, as shown in Fig. 3; however, the main effect of listener group was not significant. The interaction was also nonsignificant. The results relating the ISIB-T to proficiency were further supported by the percentage of people who demonstrated a positive ISIB-T difference at each proficiency level: LP: 61.5%; MP: 42.1%; HP: 30%.

Correlation analysis also confirmed the gradient effect of L2 proficiency. Across all Mandarin listeners, accuracies on NE speech were positively correlated with accuracies on M-US speech ($r = .653$, $p < .001$); however, as shown in Fig. 4, accuracies on NE speech were negatively correlated with the magnitude of the ISIB-T ($r = -.485$, $p < .001$).

3.2.2. Goal 3: To assess the effect of the language environment on the ISIB—

Our results indicated a divergence between the two Mandarin listener groups in the ISIB-T, but not the ISIB-L. The second set of analyses on learner-related variables among Mandarin-native listeners was carried out to address the third goal to assess the role of the language environment on the ISIB and specifically, to determine whether any influence of environment could be further specified in terms of differences in daily language exposure and use.

We collected self-reports of the frequency of L1 and L2 use on a daily basis (hours per day) including both native English listening and Mandarin-accented English listening. In addition, participants estimated their hours of English listening, English speaking, Mandarin listening, and Mandarin speaking in a typical day (Table 2). Between the M-US and M-BJ groups, t -tests revealed significant differences between the two groups in their daily use of English and exposure to English speech, and also in listening to and speaking Mandarin. On average, members of the M-US group reported listening to more English and speaking more English than the M-BJ speakers reported. Likewise, the M-US group members reported using Mandarin less in both speaking and listening than members of the M-BJ group reported. All group differences in habitual language use were significant at the .01 significance level. An unexpected outcome was a negative correlation between length of residence in an English environment and hours of English listening in the M-US group ($r = -.421$, $p < .05$). This may reflect increasing participation over time in a Chinese-speaking community as the M-US listeners settled down in the US. However, there was no correlation between the magnitude of the ISIB-T and any measure of language use. Likewise, measures of language use were uncorrelated with the magnitude of the ISIB-L. (The ISIB-L was

estimated by subtracting each Mandarin listener's accuracy on M-US speech from the average English listeners' accuracy on M-US speech.)

3.3. Goal 4: To link acoustic characteristics of native speech and foreign-accented speech with the ISIB

The last goal of the study was to explore whether plausible explanations for the ISIB could be found in the acoustic markers of stop voicing provided by native English and Mandarin-accented speakers and attention to the markers by the different groups of listeners. We conducted the following analyses to address this issue.

3.3.1. Acoustic analyses—Three acoustic measures were made of the nonword-final stops produced by the two speakers: preceding vowel duration (VD), final consonant closure duration (C2 Closure) and voicing in closure for final consonant (C2 Voicing). These measures were selected because, for stop consonants, the voicing difference was a key feature, and it has been identified as a main source of misidentification for Mandarin speakers (Bent et al., 2008; Hayes-Harb et al., 2008). A sampling of our data suggested that most incorrect responses (approximately 94%) were voicing errors rather than manner (approximately 4%) or place (less than 2%) errors. This pattern was present for all listener groups. The acoustic analyses were conducted by the first author in Praat based on waveforms and spectrograms of each nonword token (see Table 3).

A mixed-design analysis of variance (ANOVA) with voicing category (voiceless vs. voiced) as a between-items factor and speaker (NE vs. M-US) as a within-items factor was conducted using each of the acoustic measures as the dependent variable. The outcome of the analyses was the same for each acoustic measure. The main effect of speaker was significant (VD: $F(1,136) = 23.995, p < .001$, partial $\eta^2 = .150$; C2 Closure: $F(1,136) = 81.381, p < .001$, partial $\eta^2 = .374$; C2 Voicing: $F(1,136) = 49.768, p < .001$, partial $\eta^2 = .268$) because durations were longer for the NE speaker than the M-US speaker on all measures, except for the VD of voiceless tokens. The effects of voicing were also significant (VD: $F(1,136) = 226.508, p < .001$, partial $\eta^2 = .625$; C2 Closure: $F(1,136) = 167.124, p < .001$, partial $\eta^2 = .551$; C2 Voicing: $F(1,136) = 67.878, p < .001$, partial $\eta^2 = .333$) because vowel duration was longer, closures shorter, and voicing in closure greater for the voiced than the unvoiced consonants. Finally, the speaker $\times$ voicing interaction was uniformly significant (VD: $F(1,136) = 81.778, p < .001$, partial $\eta^2 = .376$; C2 Closure: $F(1,136) = 29.288, p < .001$, partial $\eta^2 = .177$; C2 Voicing: $F(1,136) = 23.656, p < .001$, partial $\eta^2 = .148$) because the NE speaker showed a larger difference on all three measures to contrast voiced vs. voiceless consonants than the M-US speaker did.

In addition, we found that, of all stimuli, only five tokens contained unreleased final stops. All of them were produced by the NE speaker, and they all led to significantly higher accuracy by NE listeners (.95 $\pm$.01) than those by M-US (.23 $\pm$.12) and M-BJ listeners (.33 $\pm$.22), $p < .001$.

3.3.2. Correlation analysis of acoustic measures and identification accuracy—

We performed correlations among the measures and the listeners' accuracies to look for patterns of association separately for the voiceless and voiced stops. The correlation matrix (see Table 4) indicated that, for voiceless word-final stop consonants, accuracy of all three listener groups was positively correlated with C2 Closure duration. This means that listeners were more likely to identify a voiceless stop as such the longer its closure duration. This is consistent with the finding that closure durations are longer for voiceless than voiced stops (e.g., Hillenbrand, Ingrisano, Smith, & Flege, 1984).

For voiced stops, the correlation matrix indicated that both NE and M-US listeners' responses were positively correlated with VD and C2 Voicing duration. In contrast, M-BJ listeners' responses were negatively correlated with C2 Closure duration. Both findings are consistent with measured differences between voiced and voiceless stops. However, the listener groups differed in the information they focus on most.

Overall, the magnitude and significance of the correlations together suggest that preceding vowel duration and C2 Voicing duration were associated with NE and M-US listeners' accuracy to identify voiced tokens, while C2 Closure duration was the feature most correlated with intelligibility of voiceless tokens. In contrast, for M-BJ listeners, C2 Closure duration was the only measure related to their performance on voicing identification.

3.3.3. Regression analyses—We conducted multiple regression analyses to further explore which of the acoustic measures are used by listeners to achieve high intelligibility scores. VD, C2 Closure and C2 Voicing duration were used to predict listeners' accuracy on each item. In the first set of analyses, accuracies were collapsed over speakers (see Table 5).

3.3.3.1. Acoustic information used for voiced vs. voiceless consonants: Regression results in Table 5 indicated that for the voiceless tokens, all three groups attended to C2 Closure. For voiced tokens, NE listeners used VD and C2 Voicing as information, while the M-BJ group used C2 Closure; M-US listeners showed sensitivity to all three measures.

3.3.3.2. Differences depending on the speaker: In a second set of regression analyses, we separated the speech tokens by speaker, and added the product of each acoustic measure and speaker group as additional predictors. Across the three listener groups, none of the regression models showed a significant result, suggesting that speaker group was generally not a moderator of the relationship between acoustic predictors and listeners' accuracy.

To summarize, voicing distinctions were found in all three acoustic features, in both NE and M-US speech, although they were more salient in the NE speech. That is to say, the native and non-native speakers did not provide different acoustic markers to distinguish voiced from voiceless coda stops, at least not along the acoustic dimensions we examined here. However, listeners did demonstrate a difference in their use of the various acoustic measures: M-BJ listeners generally neglect variability in preceding vowel duration and C2 Voicing as information for voicing, and focus solely on C2 Closure duration; native-English listeners rely on durational changes in vowel duration and C2 Voicing for voiced tokens; M-US listeners are sensitive to all three measures.

4. Discussion

4.1. Goal 1: To confirm that the ISIB-L and ISIB-T are independent phenomena

The first aim of our research was to clarify whether the interlanguage speech intelligibility benefits for talkers and for listeners are indeed distinct phenomena, as suggested by Hayes-Harb et al. (2008). Our results provided further evidence for the claim: Both the M-US and the M-BJ listeners performed significantly better than the NE listeners on M-US speech, confirming the prediction of an ISIB-L. In contrast, while the M-US listeners' performance did not differ significantly depending on whether the speaker was Mandarin-accented or not, M-BJ listeners were significantly more accurate with M-US speech than with NE speech. Thus the ISIB-T was only observed with the M-BJ group. The finding paralleled the studies by van Wijngaarden et al. (2002) and Bent and Bradlow (2003) who also found an ISIB-T. The divergent results for the two Mandarin-speaking listener groups, well-matched in age of acquisition of English and proficiency, indicate that the ISIB-L and ISIB-T do not always go

hand in hand. This implies some independence of each type of interlanguage benefit for second language users.

4.2. Goals 2 and 3: To explore the effects of proficiency and language environment on the ISIB

Our second goal was to investigate the role of L2 proficiency in speech perception, especially in contributing to the ISIB. Our data quantitatively demonstrated for the first time that the ISIB-T is gradient for English L2 (in this case L1 Mandarin) listeners with its magnitude depending on individuals' English proficiency, regardless of the language environment. In particular, the lower a listener's proficiency was, the larger was the magnitude of the ISIB-T. This finding replicated and quantified the previous finding that an ISIB-T is present only in listeners with relatively low proficiency (van Wijngaarden et al., 2002). It is important to note that, although we had only one non-native speaker who was highly proficient in English based on standardized tests and her language experience, the current result supports Bent and Bradlow's (2003) finding that the ISIB-T holds for talkers with high L2 proficiency.

The third goal of our study was to explore the influence of the language environment on identification of speech in L2 by testing two groups of participants, one in the US and one in China, who started L2 acquisition around the same time, and acquired English primarily through classroom learning. Monolingual English speakers had difficulty identifying Mandarin-accented nonwords; in contrast, L2 non-immersion learners identified speech better if it was produced by someone with whom they share a common first language. Interestingly, L2 immersion listeners were intermediate between the native-English and L2 non-immersion groups; they performed equally well with L1-accented speech and native-L2 speech as reflected by their identification scores. Moreover, while they still were less accurate than native speakers in perceiving native English speech, they did not perform worse on Mandarin-accented-speech than their non-immersion counterparts.

Our further attempt to pin down the factors in the language environment that contribute to the ISIB was partially successful. On the one hand, the M-US and M-BJ groups did differ in their daily use of L1 and L2. On the other hand, that difference was not significantly correlated with the magnitude of the ISIB-T or ISIB-L. The lack of a direct relationship between language use and the ISIB-T may be due to large individual variation in the sample. Another possible explanation is that the observed effect of environment may be the result of not only a cumulative effect of habitual use of both L1 and L2, but also the social settings in which they use L1 and L2. We collected data through self-reports to learn about the general pattern of language usage in M-US and M-BJ groups. As shown in Table 2, M-BJ group reported more passive listening than active speaking, and their exposure to native English was mostly from movies or TV shows. In contrast, M-US group spent about equal time speaking and listening, indicating that they experienced English use in conversational settings. Small differences in the kind of language input and output can possibly result in a qualitatively different behavioral pattern. Previous studies have suggested that relative usage of L1 and L2 (Flege & MacKay, 2004) and relative quantity and quality of input from native L2 speakers (Jia & Aaronson, 2003) dynamically interact with learners' developmental level and can explain age-related differences in L2 learning. Our result implies that the influence of these factors is not limited to early language acquisition but is effective in adulthood as well. After some time in a native environment (in this case, ranging from 3 to 66 months with a median of 15 months), adults in the M-US group attuned to native speech without losing sensitivity to the characteristics of Mandarin-accented-speech. This adaptive attunement helps to explain the lack of positive evidence of an ISIB-T in the M-US group. This result is consistent with Hayes-Harb et al.'s (2008) findings, and it reconciles contradictory findings from previous studies. Non-native listeners who are exposed

primarily or entirely to English spoken with an accent in their own language perceive their fellows' speech better than that of native English speakers. In contrast, it is possible that, like our M-US participants, participants who have already resided in the US for some time, may not show an ISIB-T. As exposure to native accents in L2 takes place, the "benefit" gained by a shared interlanguage diminishes.

4.3. Goal 4: To link acoustic characteristics of native speech and foreign-accented speech with the ISIB

Our attempts to pin down the source of the ISIB in the acoustic manifestations of our two speakers' stop consonants led to some interesting findings. We had just one speaker of each language, and so we cannot be certain that our analyses are general to native and Mandarin-accented English speakers. However, in agreement with previous studies (e.g., Bent et al., 2008; Flege, 1987, 1989), our acoustic analyses revealed that the Mandarin speaker produced a smaller (albeit significant) difference between voiced and voiceless word-final stops than the native English speaker on every measure. This difference in production is probably grounded in the L1 phonology of the M-US speaker in which there are no voiced stops and there are no oral stops at all in coda position.

The Mandarin listeners demonstrated much higher accuracy and more native-like perceptual responses in the voiceless context than in the voiced context. Perhaps the voiceless context was easier for listeners because Mandarin stops are always voiceless.

The existence of an ISIB-L led us to ask whether acoustic information is weighted by NE and Mandarin listeners differentially. In particular, the M-US speaker might have provided acoustic information that misled NE listeners. In addition, or instead Mandarin listeners might be more sensitive than NE listeners to the non-native acoustic patterning provided by the M-US speaker. As to the first issue, our acoustic analyses showed that the M-US speaker provided more subtle voicing differences although she did distinguish the voicing category on all three measures. Accordingly, her productions should not have misled native listeners, but perhaps made the distinction more difficult to detect. As to the second issue, we did find evidence that the three groups of listeners attended differently to the acoustic properties that we measured. The correlation and regression analyses suggested that NE listeners consistently attended to vowel duration differences (e.g., Crowther & Mann, 1992; Flege, 1989) while M-BJ listeners relied more on C2 Closure duration. Interestingly, M-US listeners once again stood between these two groups in being sensitive to all three acoustic measures. The use of a combination of various kinds of information demonstrated by this group is consistent with previous findings (Bent et al., 2008).

It is not obvious why our Mandarin listeners attended more to C2 Closure than English listeners. To our knowledge, there is no published evidence on Mandarin that suggests a closure duration difference between aspirated and unaspirated unvoiced stops (that might be detectable in non-initial position in a sentence in that language). However, there is other evidence that Mandarin listeners do not attend to preceding vowel duration and do attend to closure duration. Flege (1989) found that, even with the preceding vowel information differentiating /d/ and /t/ present, removing final release bursts impaired Mandarin listeners' identification of word-final /t/-/d/ contrasts in English words. Mandarin listeners' inattention to vowel duration and corresponding attention to closure duration present special difficulties for Mandarin listeners, because American English final stops are sometimes not audibly released (Bent et al., 2008; Jurafsky, Bell, Gregory, & Raymond, 2001). That said, future research is required to measure closure durations of stops in Mandarin and to investigate whether a difference along this dimension exists between stop consonants produced by Mandarin speakers that might explain Mandarin listeners' attention to this acoustic dimension.

One unique feature of the current study is that we directly compared late L2 listeners with or without immersion learning experience in assessing the ISIB. We found that a short period (approximately a year and half on average) of native language exposure was sufficient to produce a different strategy for extracting information from native speech, even for adult L2 listeners. The regression analyses suggested that English L2 learners in the US underwent a change in how they extract information from native speech so that they began to perceive English speech in a more native-like way. Specifically, for NE and M-US, but not the M-BJ listeners, vowel duration and voicing during closure were correlated with higher intelligibility for voiced tokens. M-US listeners were also sensitive to closure duration in voiceless tokens. Such flexibility was especially demonstrated by our M-US group members who were more likely to use both languages in daily life.

5. Conclusion

The present study suggests that the ISIB for listeners and for talkers are distinct phenomena and that the presence of an ISIB-L or ISIB-T is influenced by the language environment in which listeners use L2. Moreover, the magnitude of the benefit for talkers is modulated in a gradient manner by the proficiency of the listener. Learned weighting of L1-specific acoustic information can be a source of perceptual errors in L2 perception, but can lead to good performance on foreign-accented speech if there is a shared interlanguage between the speaker and listener. Here, we controlled for age of acquisition and English proficiency among two groups of Mandarin listeners, and our results suggest that differential language use and the quality of language input are underlying factors in effects of the language environment.

Consistent with that conclusion we found that native English and Mandarin listeners attended differently to acoustic information in the word-final consonant voicing distinction. Specifically, the most salient information extracted by native English listeners to indicate stop voicing was preceding vowel duration, whereas non-immersion Mandarin listeners were attuned to acoustic differences in closure duration. Mandarin listeners who acquired English in China but have resided in the US for some time showed sensitivity to a combination of the acoustic information for stop voicing. As a result, relatively non-native English productions as produced by the M-US speaker were still perceptually intelligible for some Mandarin listeners while they frequently prevented native listeners from identifying the targeted coda stop consonant.

Future research is required to investigate further the relationships between perceptual difficulty and closure duration in connected speech, and other specific acoustic parameters not explored here, such as amplitude information, formant transitions, burst intensity, etc. This will help determine the acoustic dimensions that contribute most to the ISIB effects in Mandarin speakers of English.

Acknowledgments

Preparation of the manuscript was supported by NICHD grant HD-001994 to Haskins Laboratories. We thank Laurie Feldman for introducing our project to collaborators in Beijing, Hua Shu for supervising collection of data and Lijuan Zou for the recruitment of participants in Beijing.

References

- Aoyama K, Flege J, Guion SG, Akahane-Yamada R, Yamada T. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*. 2004; 32(2):233–250.

- Bent T, Bradlow AR. The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*. 2003; 114:1600–1610. [PubMed: 14514213]
- Bent T, Bradlow AR, Smith BL. Production and perception of temporal patterns in native and non-native speech. *Phonetica*. 2008; 65(3):131–147. [PubMed: 18679042]
- Bolton K. Chinese Englishes: From Canton jargon to global English. *World Englishes*. 2002; 21(2): 181–199.
- Bradlow AR, Bent T. Perceptual adaptation to non-native speech. *Cognition*. 2008; 106(2):707–729. [PubMed: 17532315]
- Crowther CS, Mann V. Native language factors affecting use of vocalic cues to final consonant voicing in English. *Journal of the Acoustical Society of America*. 1992; 92(2, Pt 1):711–722. [PubMed: 1506526]
- Flege, J. Language contact in bilingualism: Phonetic system interactions. In: Cole, J.; Hualde, J., editors. *Laboratory phonology*. Vol. 9. Berlin: Mouton de Gruyter; 2007. p. 353–380.
- Flege J, MacKay IA. Perceiving vowels in a second language. *Studies in Second Language Acquisition*. 2004; 26(1):1–34.
- Flege J, MacKay IA, Meador D. Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America*. 1999; 106(5):2973–2987. [PubMed: 10573909]
- Flege JE. The development of skill in producing word-final English stops. *Journal of the Acoustical Society of America*. 1987; 82(2):433–447. [PubMed: 3624648]
- Flege JE. Chinese subjects' perception of the word-final English /t-/d/ contrast: Before and after training. *Journal of the Acoustical Society of America*. 1989; 86:1684–1697. [PubMed: 2808918]
- Hanulíková A, Weber A. Sink positive: Linguistic experience with the substitutions influences nonnative word recognition. *Attention, Perception, & Psychophysics*. 2012; 74(3):613–629.
- Hayes-Harb R, Smith BL, Bent T, Bradlow AR. The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics*. 2008; 36(4):664–679. [PubMed: 19606271]
- Hillenbrand J, Ingrisano DR, Smith BL, Flege JE. Perception of the voiced-voiceless contrast in syllable-final stops. *Journal of the Acoustical Society of America*. 1984; 76:18–26. [PubMed: 6747105]
- Imai S, Walley AC, Flege JE. Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Journal of the Acoustical Society of America*. 2005; 117:896–907. [PubMed: 15759709]
- Iverson P, Kuhl PK, Akahane-Yamada R, Diesch E, Tohkura Y, Kettermann A, et al. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*. 2003; 87(1):B47–B57. [PubMed: 12499111]
- Jia G, Aaronson D. A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*. 2003; 24:131–161.
- Jongman, A.; Wade, T.; Sereno, J. On improving the perception of foreign-accented speech. In: Sole, MJ.; Recasens, D.; Romero, J., editors. *Proceedings of the 15th international congress of phonetic sciences*. 2003. p. 1561–1564.
- Jurafsky, D.; Bell, A.; Gregory, M.; Raymond, WD. Probabilistic relations between words: Evidence from reduction in lexical production. In: Bybee, J.; Hopper, P., editors. *Frequency and the emergence of linguistic structure*. Philadelphia: John Benjamins; 2001. p. 229–254.
- Leikin M, Ibrahim R, Eviatar Z, Sapir S. Listening with an accent: Speech perception in a second language by late bilinguals. *Journal of Psycholinguistic Research*. 2009; 38(5):447–457. [PubMed: 19259819]
- MacKain KS, Best CT, Strange W. Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*. 1981; 2:369–390.
- Major RC, Fitzmaurice SM, Bunta F, Balasubramanian C. The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly*. 2002; 36:173–190.
- Munro MJ, Derwing TM. Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning*. 1995; 45:73–97.

- Munro MJ, Derwing TM, Morton SL. The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*. 2006; 28:111–131.
- Pinet M, Iverson P. Talker-listener accent interactions in speech-in-noise recognition: Effects of prosodic manipulation as a function of language experience. *Journal of the Acoustical Society of America*. 2010; 128(3):1357–1365. [PubMed: 20815470]
- Pinet M, Iverson P, Huckvale M. Second-language experience and speech-in-noise recognition: Effects of talker-listener accent similarity. *Journal of the Acoustical Society of America*. 2011; 130(3): 1653–1662. [PubMed: 21895102]
- Sancier ML, Fowler CA. Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*. 1997; 25(4):421–436.
- Smith LE, Rafiqzad K. English for cross-cultural communication: The question of intelligibility. *TESOL Quarterly*. 1979; 13:371–380.
- Stibbard RM, Lee JI. Evidence against the mismatched interlanguage intelligibility benefit hypothesis. *Journal of the Acoustical Society of America*. 2006; 120:433–442. [PubMed: 16875239]
- Van Engen KJ, Baese-Berk M, Baker RE, Choi A, Kim M, Bradlow AR. The Wildcat Corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language and Speech*. 2010; 53(4):510–540. [PubMed: 21313992]
- van Wijngaarden SJ, Steeneken HM, Houtgas T. Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America*. 2002; 111(4):1906–1916. [PubMed: 12002873]
- Zhang Y, Elder C. Judgments of oral proficiency by non-native and native English speaking teacher raters: Competing or complementary constructs? *Language Testing*. 2011; 28(1):31–50.

Fig. 1.
Two types of interlanguage intelligibility benefit.

Fig. 2.
The mean identification accuracy for stop consonants in word-final position. The error bars bracket ± 1 SE.

Fig. 3. Mean difference between accuracy with M-US speech and NE speech. A positive value indicated better performance with M-US speech. Standard errors are marked.

Fig. 4.
Relation between overall accuracy with NE speech and the magnitude of the ISIB.

Table 1

Demographic information for Mandarin listener group in Beijing, China (M-BJ) and in the US (M-US), averaged across participants.

Listener group	Age	AoA	TOEFL (total)	TOEFL (speaking)	TOEFL (listening)
M-BJ	23.5 (1.8)	10.7 (2.2)	99 (1.9)	21 (1.6)	24 (2.7)
M-US	22.6 (5.4)	9.8 (2.6)	98 (7.3)	22 (2.4)	25 (3.1)

One standard deviation presented in parentheses.

Note: AoA stands for “Age of Acquisition”.

Table 2

L1 and L2 use frequency (h/day).

Listener group	<u>English</u>		<u>Mandarin</u>	
	Listening	Speaking	Listening	Speaking
M-BJ	2.05 (1.96)	.51 (.68)	9.42 (4.74)	8.83 (5.23)
M-US	3.82 (2.05)	2.56 (1.86)	4.71 (3.76)	5.13 (3.57)

One standard deviation presented in parentheses.

Note: Exposure to native English and Mandarin-accented English are collapsed and listed under “English listening and speaking”.

Table 3

Mean acoustic measures in ms, by speaker and consonant, averaged across items.

Speaker	Target consonant	VD	C2 Closure	C2 Voicing
NE	Voiceless	187 (42)	184 (50)	12 (15)
	Voiced	372 (64)	101 (30)	39 (23)
M-US	Voiceless	209 (65)	115 (29)	7 (12)
	Voiced	300 (68)	83 (30)	15 (16)

One standard deviation presented in parentheses.

Table 4

Correlations between identification accuracy and acoustic measures.

<u>Acoustic measure</u>		<u>Accuracy</u>		
		NE	M-US	M-BJ
Voiceless	VD	-.159	-.096	-.120
	C2 Closure	.234 **	.249 **	.241 **
	C2 Voicing	.013	.073	-.017
Voiced	VD	.249 **	.238 **	.025
	C2 Closure	-.109	-.111	-.253 **
	C2 Voicing	.233 **	.268 **	.043

* and ** represent a significant correlation at the .05 and .01 significance level (2-tailed), respectively.

Table 5

Regression on acoustic measures.

Voicing	Listeners	R^2	Adjusted R^2	VD (Beta)	C2 Closure (Beta)	C2 Voicing (Beta)
Voiceless	NE	.069 [*]	.048 [*]	-.117	.216 [*]	-.024
	M-US	.065 [*]	.044 [*]	-.052	.233 ^{**}	.031
	M-BJ	.067 [*]	.046 [*]	-.074	.237 ^{**}	-.059
Voiced	NE	.122 ^{**}	.103 ^{**}	.188 [*]	-.162	.248 ^{**}
	M-US	.137 ^{**}	.118 ^{**}	.168 [*]	-.178 [*]	.291 ^{**}
	M-BJ	.08 ^{**}	.06 ^{**}	-.031	-.296 ^{**}	.136

* and ** represent significant results at the .05 and the .01 significance level (2-tailed), respectively.