LISTENING WITH A FOREIGN-ACCENT: THE ISIB EFFECT IN MANDARIN SPEAKERS OF ENGLISH

Xin Xie, Department of Psychology, University of Connecticut
Carlos A. Fowler, Department of Psychology, University of Connecticut

INTRODUCTION

Foregut-speech is a source of variability that influences the intelligibility of speech. Non-native listeners can exploit such variability and exhibit an interlanguage speech intelligibility benefit (ISIB-T). L2 learners identify foregut-speech with greater accuracy than they identify native speech (Bent & Bradlow, 2003; Stibbard & Lee, 2006). Two types of ISIB:

1) an ISIB for listeners, or an ISIB-T (Hayes-Harb et al., 2008), a benefit for non-native listeners over native listeners when speech is non-native
2) an ISIB for talkers, or an ISIB-T (Bent and Bradlow, 2003), a benefit in intelligibility for non-native listeners when speakers are non-native versus native.

QUESTION 1: Are the ISIB-T and ISIB-T independent phenomena?

The role of L2 proficiency and language environment in modulating ISIB-T

Some evidence demonstrated that ISIB-T were only present among low-proficient L2 listeners (Wargnegan et al., 2002), suggesting a role of listeners' L2 proficiency. Others failed to find any ISIB-T regardless of listeners' L2 proficiency (Hayes-Harb et al., 2008; Stibbard & Lee, 2006).

Ambient language environment may affect L2 perception, but is confounded with L2 proficiency (Pine et al., 2011).

QUESTION 2: Does L2 proficiency modulate the ISIB-T? After controlling for L2 proficiency, would language environment further affect the ISIB-T?

Acoustic characteristics of native speech and foreign-accented speech

Different acoustic information is used in systematically different manners by native and non-native listeners to differentiate contrasting phonetic segments (e.g., Flege, 1989).

QUESTION 3: Can differences in productions of stop voicing by English and Mandarin speakers and/or these listener groups' differential attention to the acoustic specifications explain the observed ISIBs?

CURRENT STUDY

We investigated the role of language proficiency and ambient language environment on the ISIB by examining the intelligibility of native and Mandarin-accented English speech for three groups of listeners: native English (NE), Mandarin-speaking Chinese listeners (M-US) in the 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. Thus, we directly compared late L2 listeners with or without immersion learning experience in assessing the ISIB-T.

METHODS

PARTICIPANTS

Undergraduate and graduate students: 34 monolingual native-English listeners (NE) 25 native-Mandarin speakers in the US (M-US) who speak English as an L2 3D native-Mandarin speakers in Beijing, China (M-BJ) who speak English as an L2

A self-report survey was used to collect information about participants' language background.

For the M-US group, the length of residence in the US ranged from 3 to 66 months (M = 19.8 months, SD = 18 months).

Nonword Transcription Task

Stimuli: 32 items of “ENGLISH” monosyllabic nonwords with stop consonants in word-final position e.g., ved, zib, sheeg, sit, doop, rook

Each list had 20 nonwords 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 presented.

Recorded by 2 age-matched speakers: 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 recorded

Design and Procedure:

- Nonword items counterbalanced in 2 lists
- Participants randomly assigned to one of the lists
- Repetition times monitored by participants

RESULT

Nonword transcription results

3 x 2 x 2 ANOVA design with between-group factor listener group (NE vs. M-US vs. M-BJ), and within-group factor speaker (NE vs. US) and voicing (voiced vs. voiceless coda stop consonant)

Main effect of listener group, F(2,86) = 3.134, p < .05

Main effect of speaker, F(1,86) = 382.443, p < .001

Main effect of voicing, F(1,86) = 326.73, p < .001

Interaction between speaker and listener group, F(2,86) = 16.773, p < .001; voicing and speaker, F(1,86) = 157.54, p < .001; three-way interaction, F(2,86) = 32.789, p < .001

ISIB-T for both Mandarin groups:

For the M-US speech, M-BJ listeners and M-US listeners > NE listeners, p < .05

For the NE speech, NE listeners > M-US listeners and M-BJ listeners, ps < .001

No difference between M-US and M-BJ listeners

ISIB-T for M-BJ groups, but not for M-US

M-BJ group: the M-US speaker was more intelligible than the NE speaker, t(29) = 2.28, p < .05

M-US group: two speakers equally intelligible, t(24) = -0.292, p = .054

NE group: more accurate with NE speech than M-US speech, t(33) = -25.122, p < .001

The role of language proficiency and language environment

The magnitude of ISIB-T: subtract accuracy on NE speech from accuracy on M-US speech

Listener's proficiency determined by their accuracy on NE speech

- Highly-proficient (HP): > .85
- Moderate-proficient (MP): [.75, .85]
- Low-proficient listeners (LP): < .75

The average accuracy of the NE listener group was 94 ± 0.4

M-US and M-BJ groups did not differ, t(21) = 2.15, p = .098

Across all Mandarin listeners,

- Accuracy on NE speech were negatively correlated with the magnitude of the ISIB-T, r = -.458, p < .001
- Accuracy on NE speech were positively correlated with accuracies on M-US speech, r = .653, p < .001

Linking acoustic characteristics of native and foreign-accented speech with the ISIB

Acoustic analysis

- Main effect of speaker: longer durations in NE speech than M-US speech, except for the VD of voiceless tokens, ps < .001
- Main effect of voicing: longer vowel duration and voicing in closure, shorter closures for the voiced than the unvoiced consonants, ps < .001
- Larger difference to contrast voiced vs. voiceless consonants in NE speech than M-US speech, ps < .001

Correlation analysis

- M-BJ listeners focus solely on C2 closure duration; NE and M-US listeners rely on durational changes in vowel duration and C2 voicing for voice token

DISCUSSION

The ISIB-T and ISIB-T are independent phenomena

- Two Mandarin-speaking listener groups, well-matched in age of acquisition of English and proficiency, outperformed native English listeners in identification of Mandarin-accented speech (showing an ISIB-L). However, the ISIB-T was only observed with the M-BJ group, suggesting a role of ambient language environment on the ISIB-T.

The divergence in M-US and M-BJ groups implies some independence of each type of ISIB for second language users.

L2 Proficiency modulates the ISIB-T

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. The lower a listener’s proficiency was, the larger was the magnitude of the ISIB-T.

Acoustic Characteristics of Native Speech and Foreign-accented Speech

The ISIB-T gradient for English L2 should not have misled native listeners, but made the distinction more difficult to detect.

Acoustic information is weighted by NE and Mandarin listeners differently. Specifically, for NE and M-US, but not the M-BJ listeners, vowel duration and voicing during closure were correlated with higher intelligibility for voiceless tokens. M-US listeners were also sensitive to closure duration in voiceless tokens. Results suggest 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.

ACKNOWLEDGEMENTS: This work was supported by NICHD grant HD-001994 to Haskins Laboratories.

WORKS CITED:


