

A PRELIMINARY STUDY ON THE INTERLANGUAGE SPEECH INTELLIGIBILITY BENEFIT FOR ENGLISH-MANDARIN BILINGUAL L2 LEARNERS

Guo Li and Peggy Mok

Department of Linguistics and Modern Languages, The Chinese University of Hong Kong
liguo.vivian@gmail.com, peggymok@cuhk.edu.hk

ABSTRACT

Previous studies into interlanguage speech intelligibility benefit (ISIB) have focused on the influence of subjects' native language (L1) on the phonetic production and perception in their second language (L2). However, no research so far has examined the effect of the listeners' exposure and training in a second language (L2) on their understanding of L2-accented native language (L1). This paper aims to address this issue with subjects whose L1 is English, L2 is Mandarin. Characteristics of Mandarin-accented English include the devoicing of word-final consonants, and the insufficient distinction of the vowel pairs /i:/ - /ɪ/ and /e/ - /æ/. These features could negatively affect listeners' understanding of contrastive word pairs. In this study, 9 native Mandarin listeners, 9 monolingual English listeners and 9 English-Mandarin bilinguals were asked to listen to recordings of Mandarin-accented English and identify minimal pairs involving the above consonant and vowel contrasts. Results show that among all three groups of subjects, native Mandarin listeners scored the highest accuracy, but English listeners with training in Mandarin and monolingual English speakers had similar scores. These findings support the existence of ISIB for Mandarin, and call for further study on bilingual L2 learners.

Index Terms— speech perception, intelligibility, Mandarin, English

1. INTRODUCTION

There are many factors influencing the intelligibility of speech for non-native listeners. One of which is called the Interlanguage Speech Intelligibility Benefit (ISIB) suggested by Bent and Bradlow [1]. They conducted an experiment asking native Chinese, Korean and English listeners and listeners with other L1 backgrounds to perform sentence recognition task, using recordings of native Chinese, Korean and English talkers reading simple English materials. They found that for non-native listeners, the talkers who shared with them the same native language

were more intelligible than or at least as intelligible as native English talkers. Bent and Bradlow attributed the ISIB to the overall shared phonetic and phonological knowledge between non-native talkers and the non-native listeners. For example, when a non-native talker produced vowels of his/her L1 (e.g. Chinese) as substitutes for vowels of L2 (e.g. English), native English listeners might be misled while non-native listeners (Chinese) would be more likely to understand the intended vowels given the shared phonological and phonetic knowledge of the Mandarin sound system.

Many studies have investigated the ISIB with mixed results. For example, Hayes-Harb et al [2] asked Mandarin listeners and native English listeners to perform in a word identification task, using minimal pairs in English with final-voicing contrasts as material. They found ISIB for Mandarin listeners, especially those with limited English proficiency. They also attributed ISIB to the shared phonological system between Mandarin listeners and talkers. On the contrary, Munro et al [3] asked Cantonese, Japanese, Mandarin and English listeners to rate the English speech produced by Cantonese, Japanese, Polish and Spanish speakers in terms of intelligibility and accentedness. They found ISIB for Japanese listeners only, but not for Cantonese listeners. Similarly, using lecture recordings in English spoken by speakers of Mandarin, Japanese, Spanish and American English, Major et al [4] found ISIB for Spanish listeners only, but not for Japanese and Mandarin listeners.

The discrepant results suggest that the mechanisms behind ISIB are complex and thus in need of further exploration. Specifically, if the presence of ISIB is attributable to the shared phonological system between the talkers and the listeners, then ISIB should be extendable to anyone who shares the same knowledge, i.e., bilingual L2 learners as well. If Mandarin listeners could understand Mandarin talkers speaking English better than native English listeners do because of the shared phonological knowledge, then native English listeners who have learned Mandarin as a second language (i.e., sharing the same Mandarin knowledge) should also understand Mandarin talkers better than monolingual native English listeners do

as well. However, so far, there is no study investigating this possibility. The current study is a preliminary investigation to examine whether ISIB found for Mandarin listeners can also be extended to English-Mandarin bilingual L2 learners.

Mandarin-accented English is characterized by these features: 1) devoicing of word-final stops; 2) insufficient contrast between /i:/ and /ɪ/; and 3) non-distinction between /ɛ/ and /æ/ [5, 6, 7]. We compared listeners' identification of minimal pairs in English involving these contrasts to evaluate the ISIB for Mandarin listeners and English-Mandarin bilingual L2 learners.

2. METHOD

2.1. Subjects

Three listener groups participated in this study: 9 native Mandarin speakers (NM); 9 native English speakers who have learned Mandarin as a second language (BI) and 9 native monolingual English speakers (NE). The 9 female NM listeners had a mean age of 21. They came from various regions in China and spoke Mandarin as their first language. They were students at CUHK. The 9 BI listeners (5 females, 4 males) with a mean age of 24 were recruited from advanced courses in Mandarin at the Yale-China Chinese Language Center at CUHK. 7 out of the 9 BI listeners were exchange students from the USA. The 9 NE listeners (6 females, 3 males) with a mean age of 25 were recruited through word of mouth in Hong Kong. One was from the UK, one was from Australia, two were from Canada, and 5 were from the USA. None of them understood Mandarin. All listeners reported no history of speech or hearing problems. They volunteered to participate in this study.

2.2. Materials

A list of 36 English monosyllabic words in CVC structure was used: deep, dip, peace, piss, sheep, ship, beat, bit, dead, dad, pet, pat, bet, bat, bed, bad, pick, pig, peck, peg, back, bag, buck, bug, cop, cob, cap, cab, cup, cub, rip, rib, bid, bud, but, bead. These 36 words form 20 minimal pairs. There are four minimal pairs for each of the five contrasts: vowels: /i:/ - /ɪ/, /ɛ/ - /æ/ and word-final stops: /p/-/b/, /t/-/d/, /k/-/g/.

Seven female native Mandarin speakers produced the 36 target words in a short carrier phrase three times. Because of time limit, only materials from three female speakers were used for the listening experiment. These three speakers were chosen because of the heavy Mandarin accent in their English. One token of each target word was taken as the materials for the word identification experiment, resulting in 120 word tokens in total (20 word pairs × 2 words × 3 speakers). The target words were excised from

the carrier phrase, i.e., only the target words would be played to the listeners for identification.

2.3. Procedure

The subjects listened to the recordings in a semi-random order over a headphone in a quiet room. Care was taken to ensure that no minimal pair would appear in a row. For each recording, the subjects were asked to choose from two given words on an answer sheet (e.g. cob, cop). They were allowed to listen to the recordings repeatedly if necessary, and most listeners took advantage of that. In addition, they were asked to give a confidence rating ranging from 0 (no confidence) to 7 (full confidence) for each answer.

Upon completion, the subjects were asked to fill in a questionnaire about their language backgrounds.

3. RESULTS

The number of correct answers of each listener was counted and converted to percentage. Table 1 shows the overall performance of the three listener groups. A one-way ANOVA confirms that the overall performance differed significantly across groups [$F(2,24) = 4.142, p = 0.028$]. Post-hoc Bonferroni comparisons show that NM listeners performed significantly better than BI listeners ($p = 0.04$) and with a trend for significance when compared with the NE listeners ($p = 0.098$), while the BI listeners and NE listeners did not differ ($p = 1.0$). The results indicate that ISIB exists for Mandarin listeners, but there was no advantage for the English-Mandarin bilingual L2 learners over the monolingual English listeners.

Table 1. Overall accuracy of the three listener groups

Listeners	MEAN	S.D.
NE	56.02%	2.67
BI	55.09%	2.90
NM	61.20%	7.90

Figures 1 and 2 on the next page show the results of the specific consonant and vowel contrasts respectively. Similar to the overall results, the NM listeners also performed the best in these individual contrasts, while there is no consistent pattern for the NE and BI listeners.

One-way ANOVAs were conducted for each contrast. Results show that the accuracy difference of the listener groups was significant for the /p/ - /b/ consonant contrast [$F(2,24) = 3.759, p = 0.038$], and approaching significance for the /i:/ - /ɪ/ vowel contrast [$F(2,24) = 2.824, p = 0.079$]. The differences of other contrasts were not significant.

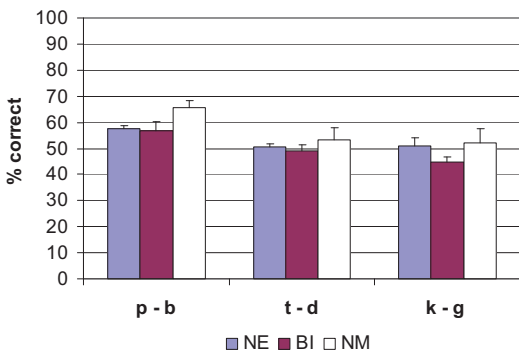


Figure 1. Accuracy of the three listener groups on word-final consonants. Error bars show one standard error.

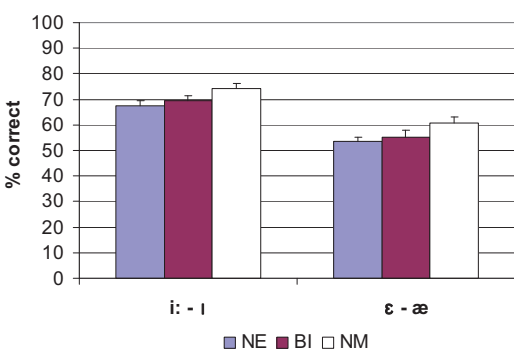


Figure 2. Accuracy of the three listener groups on vowel contrasts. Error bars show one standard error.

It is worth noting that among the same type of contrasts (i.e., word-final consonant or vowel), specific contrasts differed in terms of ease of identification for all listener groups. The /p/ - /b/ contrast was the easiest among the word-final stops. In fact, the accuracy of the /t/ - /d/ and /k/ - /g/ contrasts just hovered at chance level (50%) even for NM listeners, which shows that these contrasts in Mandarin-accented English were really difficult to discriminate. Further acoustic analysis on the listening materials is needed to confirm why there was such a difference for the consonants. As for vowels, the /i:/ - /ɪ/ contrast was much easier than the /ɛ/ - /æ/ contrast. This echoes well with previous studies showing that while the /i:/ - /ɪ/ contrast was only insufficiently distinguished in Mandarin-accented English, there was almost no distinction for the /ɛ/ - /æ/ contrast in Mandarin-accented English.

The listeners were also asked to indicate their confidence level for each answer, with 0 being no confidence at all and 7 showing full confidence. Table 2 shows the overall confidence rating of the three listener groups. We can see that the ratings of the three listener groups were quite similar, with NM listeners having slightly higher rating, followed by NE and BI listeners. The same

pattern can be observed in the confidence ratings for individual contrasts. One-way ANOVAs were conducted on both the overall rating and ratings for individual contrasts, but none of the comparison was significant. One interesting finding of the confidence ratings is that, although the rating differences were not significant, the standard deviations for the NM listener group were larger than the BI and the NE listeners for both overall and individual contrasts. This indicates that there was much individual variation among the NM listeners. We examined the ratings of individual NM listeners and found that 3 out of the 9 NM listeners consistently gave very low ratings (2 or 3 out of a 7-point scale) while the other 6 NM listeners gave much higher ratings. However, the accuracy of the 3 listeners with low confidence was actually quite comparable with other NM listeners. This interesting situation and the similarity of confidence ratings among the three listener groups indicate that performance and confidence do not necessarily correlate.

Table 2. Overall confidence rating of the listener groups

Listeners	MEAN	S.D.
NE	4.48	0.46
BI	4.18	1.03
NM	4.51	1.57

4. DISCUSSION

Our results indicate that native Mandarin listeners could understand the English minimal pairs produced by native Mandarin speakers better than native English and English-Mandarin bilingual listeners did. This is consistent with the findings in Bent and Bradlow [1] and Hayes-Harb et al [2], which indicates that ISIB does exist for Mandarin.

Nevertheless, such ISIB was not found across all consonant and vowel contrasts, which was not explored by previous studies before. Previous studies on ISIB adopted a broader approach and did not focus on specific consonant and vowel contrasts. For example, Bent and Bradlow [1] checked listeners' transcriptions of simple sentences and counted the correct transcription of key words. Major et al [8] examined ISIB by testing listeners' comprehension of short lectures. As pointed out by Hayes-Harb et al. [2], when listeners were presented with sentences, a mixture of factors might be at play, some of which may be irrelevant to ISIB. Hayes-Harb et al. [2], Bent et al. [5] and Smith et al. [8] used isolated words as stimuli, and studied word-final stops but none examined specific consonant contrasts and none included any vowel contrasts on ISIB. Our study incorporated both types of segmental contrasts in the experiment and has shown that even within the same segmental type, intelligibility could also differ. Sharing the same phonetic and phonological system between the speakers and listeners does not guarantee intelligibility in an

across-the-board manner. It is particularly revealing to find that even NM listeners could perform at chance level for some contrasts. Future studies on ISIB should investigate which contrasts are more conducive to ISIB and which are not, and whether ISIB is language-specific.

The main research question of this study is whether ISIB can be extended to bilingual L2 learners who also share the phonological system which gave rise to ISIB. Our results show that BI listeners performed similarly with NE listeners who did not know Mandarin at all, which gave a negative answer to the research question. However, before we can firmly reject this hypothesis, there are some factors which we should re-consider. First, although the BI listeners had plenty of exposure to Mandarin, they did not have much exposure to Mandarin-accented English. Being native English speakers, they might have relied more on their native English knowledge to distinguish the minimal pairs in the experiment, rather than transferring their learned Mandarin knowledge to understand Mandarin-accented English. Second, the proficiency level of the bilingual learners may also play a role. The BI listeners may need to have an even higher Mandarin proficiency level in order to benefit from the ISIB. Third, the sample size of the listener groups ($n = 9$) may be too small to yield significant differences. This was further aggravated by the fact that there was more individual variation in the BI group than the NE group, as evident in the larger standard deviations in both accuracy (Table 1) and confidence level (Table 2) of the BI group. Whatever the reason, our results indicate that simply sharing a phonological system may not be enough to give rise to ISIB. More studies with bilingual L2 learners are needed to explore the complex mechanisms behind ISIB.

Finally, it will be interesting to further investigate Mismatched Interlanguage Speech Intelligibility Benefit (MISIB) for Cantonese speakers/listeners. MISIB refers to the ISIB found when non-native listeners and non-native talkers do not share the same native language. A good example of MISIB can be found in Bent and Bradlow [1] where Mandarin listeners found high-proficiency Korean talkers more intelligible than native English talkers despite their difference in LI. As the features of the consonant and vowel contrasts of Mandarin-accented English used in this study are also found in Hong Kong English [9], it is quite possible that MISIB can be found for Cantonese listeners, but so far no study has investigated this possibility. Moreover, Mok, Setter and Low [10] have shown that juncture pairs in Hong Kong English are more intelligible than those in Singapore and British English. It will be useful to compare intelligibility of Mandarin-accented English and Hong Kong English with different listener groups.

To conclude, our preliminary study has confirmed the existence of ISIB for Mandarin as shown in previous studies. We have also shown that individual segmental contrasts can differ in intelligibility. Our results call for

further investigation of whether ISIB can be extended to situations where the talker and listeners do not match in terms of their native language, but share the same phonological system through L2 learning. Possible factors that can affect ISIB for bilingual L2 learners were discussed. It is hoped that our study has shed some new light on the interesting topic of intelligibility of non-native speech.

5. ACKNOWLEDGEMENTS

We would like to thank Prof. Gloria Zhang, Ms. Carmen Chiu, and the Yale-China Chinese Language Center at CUHK for facilitating access to English-Mandarin bilingual participants in this study. We are also grateful to Miss Li Jingwen, Miss Zuo Donghui, Miss Huang Cong, Miss Wei Jin, Miss Jiang Yuchen, Miss Zhang Xinyun and Miss Li Lewei who assisted with recording. We thank the 27 listeners in the experiment for their voluntary participation.

6. REFERENCES

- [1] T. Bent, and A.R. Bradlow, "The interlanguage speech intelligibility benefit," *Journal of the Acoustical Society of America*, 114, 1600-1610, 2003.
- [2] R. Hayes-Harb, B.L. Smith, T. Bent, and A.R. Bradlow, "The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts," *Journal of Phonetics*, 25, 437-470, 2008.
- [3] M.J. Munro, T.M. Derwing, and S.L. Morton, "The mutual intelligibility of L2 speech," *Studies in Second Language Acquisition*, 28, 111-131, 2006.
- [4] R.C. Major, S.F. Fitzmaurice, F. Bunta, and C. Balasubramanian, "The effects of nonnative accents on listening comprehension: implications for ESL assessment," *TESOL Quarterly*, 36, 173-190, 2002.
- [5] T. Bent, A.R. Bradlow, and B.L. Smith, "Production and perception of temporal patterns in native and non-native speech," *Phonetica*, 65, 131-147, 2008.
- [6] J.E. Flege, M.J. Munro, and L. Skelton, "Production of the word-final English /t/ - /d/ contrast by native speakers of English, Mandarin and Spanish," *Journal of the Acoustical Society of America*, 92: 128-143, 1992.
- [7] J.E. Flege, and O. Bohn, "Effects of experience on non-native speakers' production and perception of English vowels," *Journal of Phonetics*, 25: 437-470, 1997.
- [8] B.L. Smith, A.R. Bradlow, and T. Bent, "Production and perception of temporal contrasts in foreign-accented English," *In Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, 519-522, 2003.
- [9] Setter, J., C.S.P. Wong, and B.H.S. Chan, *Hong Kong English*, Edinburgh, Edinburgh University Press, 2010.
- [10] P. Mok, J. Setter, and E.L. Low, "The perception of word juncture characteristics in three varieties of English," *In Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS)*, 1410-1413, 2011.