

# **Interlanguage Speech Intelligibility Benefit and the Mental Representation of Second Language Speech Sounds**

Patrick Chun Kau Chu<sup>1</sup> and Marcus Taft<sup>1</sup>

<sup>1</sup>*University of New South Wales, Sydney, Australia <pchu@psy.unsw.edu.au>*

**This study investigates the perception of Cantonese-accented and native English speech by native Cantonese and English listeners. Listeners transcribed monosyllabic word pairs which are expected to be confused and mispronounced by Cantonese speakers in two separate experiments (e.g. ‘thin’ in Experiment 1 and ‘fin’ in Experiment 2). Results show that Cantonese listeners have an advantage over native listeners in understanding Cantonese-accented speech for words that are treated as homophones due to first language transfer (e.g. ‘thin’ - ‘fin’). Our study also demonstrates that Cantonese speakers can perceive and produce second language speech contrasts, though in a non-native manner. Issues related to the mental representation of second language speech sounds will be discussed.**

## INTRODUCTION

Two kinds of interlanguage speech intelligibility benefits (ISIB) have been mentioned in which non-native speakers may have an advantage over native speakers in understanding accented speech (Hayes-Harb et al., 2008). The first type, ISIB-Talker (ISIB-T), suggests that non-native speech is more intelligible than native speech to non-native listeners who share the same language background as the non-native speaker. The second type, ISIB-Listener (ISIB-L), suggests that non-native listeners have an advantage over native listeners in understanding non-native speech spoken by speakers from the same language background. ISIB was first proposed by Bent and Bradlow (2003) where they found that non-native speech from high proficiency talkers is as intelligible as native speech to non-native listeners. However, equal performance in understanding accented speech is not actually an intelligibility benefit. Several subsequent studies (e.g., Hayes-Harb et al.; 2008, Smith et al., 2009) found that non-native listeners understood native speech better than speech produced in their own accent. Concerning ISIB-L, the results are inconsistent among different studies and the effect can only be observed with dense neighborhood words (Imai et al., 2004), particular language groups (Munro et al., 2006) and low proficiency speakers and listeners (Bent et al., 2008; Hayes-Harb et al., 2008).

The inconsistent findings for ISIB-L may also be due to the words that are included in the test, particularly the inclusion of words where there are phonological errors

when pronounced by non-native speakers and the mispronounced word becomes another word. For example, Cantonese speakers may mispronounce the word ‘thin’ as ‘fin’ due to the lack of the phonemic category /θ/ in their first language (L1). Native English listeners will probably interpret this word as ‘fin’ while Cantonese listeners may interpret it as either ‘thin’ or ‘fin’ if they assimilate the English phoneme /θ/ to their L1 phoneme /f/. If the word frequency of the intended word is higher than that of the mispronounced word, as in the case of ‘thin’ being pronounced ‘fin’, it is more likely that Cantonese listeners can correctly interpret the word as ‘thin’. On the other hand, if the word frequency of the intended word is lower than that of the mispronounced word, as in the case of ‘buzz’ being pronounced like ‘bus’ by Cantonese speakers, then it is more likely that Cantonese listeners will misinterpret the word as ‘bus’ in the same way that native English listeners will. We hypothesize that ISIB-L is caused by the items where the intended words have a higher word frequency than the mispronounced word.

Within this account, we assume that the cause of foreign accent is second language (L2) learners’ inability to perceive L2 phonemic contrasts. As a result, L2 phonemes are assimilated into L1 categories. For example, the English phoneme /θ/ is assimilated to Cantonese phoneme /f/ and the words ‘thin’ and ‘fin’ are therefore homophones for Cantonese listeners. The relative high frequency word ‘thin’ will be activated first when they hear either /θɪn/ or /fɪn/. In other words, there is a single phonological system for the two languages. An alternative hypothesis is that listeners can perceive the phonemic contrasts in the L2 that are absent in their L1 and foreign accents are due to articulatory difficulty only. This suggests that there are two separate phonological systems for the two languages. The second experiment will address this issue.

## EXPERIMENT 1

The first experiment examines whether ISIB-T and ISIB-L can be found among Cantonese speakers and listeners and identifies whether the cause of ISIB-L is due to relative word frequency. Twenty listeners each of Cantonese and English background were asked to write down the English words that they heard in a monosyllabic word transcription task. The words were recorded by a Cantonese speaker with a strong accent and a native Australian English speaker.

The design of the study is shown in Table 1. In the high frequency baseword (HF) condition, the 20 intended words (baseword) have a higher word frequency than the mispronounced word (e.g. ‘thin’→‘fin’). In the low frequency baseword (LF) condition, the 20 intended words have a lower frequency than the mispronounced

word (e.g. ‘buzz’→‘bus’). Unambiguous items consist of 40 words in which Cantonese speakers do not produce any phonological errors (e.g. ‘low’). All listeners transcribed 40 words spoken by the Cantonese speaker followed by 40 words spoken by an English speaker.

**Table 1**  
**Experimental conditions for Experiments 1 and 2**

Native language of the speaker	Experiment 1		Experiment 2	
	Cantonese	English	Cantonese	English
Unambiguous items	‘low’ /loʊ/	‘low’ /loʊ/	‘low’ /loʊ/	‘low’ /loʊ/
High frequency baseword (HF)	‘thin’ /fɪn/	‘thin’ /θɪn/	‘fin’ /fɪn/	‘fin’ /fɪn/
Low frequency baseword (LF)	‘buzz’ /bʌz/	‘buzz’ /bʌz/	‘bus’ /bʌs/	‘bus’ /bʌs/

**Table 2**  
**Mean percentage of correct response in Experiment 1**

	Speakers					
	Unambiguous		HF		LF	
	Cantonese	English	Cantonese	English	Cantonese	English
Cantonese listeners	75	81	54	73	32	62
English listeners	72	95	24	90	25	96

The mean percentage of correct responses for each condition, speaker and listener groups are shown in Table 2. The main effects and all two and three way interactions effect including the listener, speaker and conditions are significant ( $p$ 's < 0.001). ISIB-T is not found as Cantonese listeners found native English speech more intelligible than Cantonese-accented speech in all conditions ( $p$  < 0.01). ISIB-L is not found for the unambiguous items as English listeners understood Cantonese-accented speech as well as Cantonese listeners ( $p$  = 0.128). However, ISIB-L is found in both the HF condition ( $p$  < 0.001) and the LF condition ( $p$  = 0.032), as Cantonese listeners understood Cantonese-accented speech better than English listeners.

The word frequency effect emerges for Cantonese listeners no matter whether they were listening to native English speech (HF: 73% vs LF: 62%) or Cantonese-accented speech (HF: 54% vs LF: 32%),  $p$ 's < 0.05. This suggests that Cantonese listeners may assimilate second language phonemes (e.g. /θ/, /z/) to their first language phonemic categories (e.g. /f/, /s/), such that ‘thin’ becomes homophonic with ‘fin’, as does ‘buzz’ with ‘bus’. When they hear these homophones they will tend to recognize it as the more frequent word, and hence a lower accuracy score is observed for the LF condition.

## EXPERIMENT 2

In this experiment, we used the word transcription task again with 20 native English listeners and Cantonese listeners and test whether they could perceive the phonemic contrasts made by native English and Cantonese speakers (see Table 1). In Experiment 1, the speakers were asked to pronounce words that are easily mispronounced by Cantonese speakers (e.g. ‘thin’ mispronounced as ‘fin’). In Experiment 2, both the Cantonese and native English speakers were asked to pronounce the mispronounced word in Experiment 1 (e.g., ‘fin’). If Cantonese listeners cannot perceive phonemic contrasts in their L2, they should show similar activation of the baseword response (e.g. ‘thin’) no matter whether Cantonese or native English speakers pronounce the word ‘thin’ or ‘fin’. The results for the HF and LF conditions are collapsed together and shown in Table 3 as there is no interaction with the conditions.

**Table 3**  
**Mean percentage of baseword responses in Experiments 1 and 2**

	Cantonese speaker		English speaker	
	Experiment 1	Experiment 2	Experiment 1	Experiment 2
Cantonese listeners	43	33	68	21
English listeners	25	22	93	2

For Cantonese listeners, there are more baseword responses in Experiment 1 than Experiment 2 ( $p < 0.001$ ), and the difference is larger for the speech produced by the native English speaker than the Cantonese speaker ( $p < 0.001$ ). This shows that Cantonese listeners can perceive the phonemic contrasts in their L2. Moreover, it suggests that there are some acoustic differences between the mispronounced word (e.g. ‘fin’) in Experiment 1 and the target word (e.g. ‘fin’) in Experiment 2 in the Cantonese-accented speech such that the former is more easily recognized as the intended word by Cantonese listeners. For English listeners, there are significantly more baseword responses in Experiment 1 than Experiment 2 for words spoken by the native English speaker ( $p < 0.001$ ). However, there is no significant difference between the percentage of baseword responses in Experiment 1 and 2 for words spoken by the Cantonese speaker ( $p = 0.244$ ). This suggests that English listeners can hear the phonemic contrasts made by native English speakers only. Although the Cantonese listeners detect acoustic differences in the Cantonese-accented speech between the mispronounced words in Experiment 1 and the target words in Experiment 2, those acoustic cues cannot be picked up by English listeners.

## DISCUSSION AND CONCLUSION

In Experiment 1, word frequency effect is observed when Cantonese listeners perceive native English and Cantonese-accented speech. This suggests that there is a single phonological system for the two languages so that words like ‘thin’ and ‘fin’ are treated as homophones for Cantonese listeners. However, in Experiment 2, we found that Cantonese speakers could perceive and produce L2 speech contrasts, though in a non-native like manner. This may suggest that there are two phonological systems for the two languages. One plausible explanation for the contradictory conclusions drawn from the two experiments is that non-native listeners are gradually developing two phonological systems from one single phonological system. This is in agreement with Flege’s (1995) Speech Learning Model in which L2 learners are developing two separate categories for the “new” sounds while “similar” sounds in the L2 are still treated as a single category. If we divide the non-native listeners into high and low phonological proficiency, we expect the latter to show a greater word frequency effect, and the former to show a greater difference between Experiment 1 and Experiment 2 in baseword responses for native English speech.

## REFERENCES

- Bent T and Bradlow AR. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, **114**, 1600-1610.
- Bent T, Bradlow AR and Smith BL. (2008). Production and perception of temporal patterns in native and non-native speech, *Phonetica*, **65**, 131-147.
- Flege J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed), *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, pp. 229-273.
- Hayes-Harb R, Smith BL, Bent T and Bradlow AR. (2008) The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts, *Journal of Phonetics*, **36**, 664-679.
- Imai S, Walley AC and Flege JE. (2004). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners, *The Journal of the Acoustical Society of America*, **117**, 896-907.
- Munro MJ, Derwing TM and Morton SL. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, **28**, 111-131.
- Smith BL, Hayes-Harb R, Bruss T and Harker A. (2009). Production and perception of voicing and devoicing in similar German and English word pairs by native speakers of German, *Journal of Phonetics*, **37**, 257-275.