PHONETIC AND LEXICAL EFFECTS ON SPEECH PERCEPTION OF FALLING TONES

Chia-Hsin Yeh & Yen-Hwei Lin

Michigan State University, USA
chiahsin@msu.edu; liny@msu.edu

ABSTRACT

This study investigates phonetic and lexical effects on speech perception of high-falling tones and low-falling tones in Hakka and Mandarin. An identification task and a lexical decision task were conducted on 30 participants for each language. The results illustrate that onset pitch and pitch drop are more crucial phonetic attributes than offset pitch for native speakers to distinguish high-falling and low-falling tones. In addition, lexical effects have been found to exert a significant influence on tonal perception. The findings suggest that a model of speech processing should consider phonetic and lexical knowledge, both low-level and high-level representations, as cognitive fundaments.

Keywords: speech perception, falling tone, lexical frequency, exemplar model, Hakka

1. INTRODUCTION

The less understood effects from prosodic contexts on pitch variations have made tonal confusion in Chinese a less studied area, except for the well-known case of T2-T3 confusion in Mandarin. Moore and Jongman [6] show that similar pitch contour between T2 and T3 is a crucial cause of the tonal confusion. Beside Mandarin T2 and T3, some other tonal contrasts that share similar phonetic features may also lead to confusion. The phonetic similarity of tonal contrasts can be derived from intrinsically similar categories (invariant relation) or prosodically conditioned similar variations (variant relation). This study focuses on the perception of high-falling and low-falling tones in two Chinese languages, Hakka (Hai-lu variety) and Mandarin. In Table 1, the number given after each tone stands for pitch height and pitch contour, with 5 indicating the highest pitch and 1 the lowest as proposed in Chao [2]. In Hakka, the two falling tones are intrinsically similar categories, as proposed by Lo [5], whereas Mandarin T3 has a low-falling variant with a falling pitch contour similar to T4 in connected speech, as described in Chao [2]. The two pairs of falling categories are spectrographically similar, but phonologically different. The goal of this study is to identify the crucial phonetic and lexical factors influencing the perception of high-falling and low-falling tones in the two languages.

Table 1: Tonal Inventories of Hakka and Mandarin. (H=high, L=low)

Types		Hakka	Mandarin
Level	H	55	55 (T1)
	L	22	n.a.
Rising		13	35 (T2)
Dipping		n.a.	214 (T3)
Falling	H	53	51 (T4)
		51 (allophone)	53 (T4 allophone)
	L	31	21 (T3 allophone)
Checked	H	5	n.a.
	L	2	n.a.

2. PERCEPTUAL FACTORS

In this study, we investigate how phonetic similarity and lexical frequency influence perceptual processing of falling tones and their variations. As Pierrehumbert [7] proposes, multiple linguistic and paralinguistic information, such as phonetic, phonological, and lexical knowledge, is stored as perceptual representation in lexical memory to facilitate speech processing. Each kind of information is a perceptual factor for spoken word recognition.

For phonetic effects, following Yeh [8], three phonetic characteristics, onset pitch height, offset pitch height, and pitch drop were manipulated for the experiment. Onset pitch refers to the initial pitch peak; offset pitch refers to the ending pitch valley; pitch drop refers to the difference between the peak and the valley. For lexical effects, lexical frequency is considered.

If these four perceptual factors are crucial to the perception of high-falling versus low-falling tones, then the hypotheses are as follows:

(1) Four research hypotheses
   a. Onset pitch perturbations will cause tonal confusion between high-falling and low-falling tones, and a perceptual boundary of the two categories will be found.

b. Offset pitch perturbations will cause confusion of the two tones, and a perceptual boundary will be found.
c. Pitch drop perturbations will cause confusion of the two tones, and a perceptual boundary will be found.
d. Frequent words will more likely be chosen, and a significant difference between responses of the frequent and less frequent words will be found.

3. METHODOLOGY

In the experimental design, the three phonetic factors are set up as variables in the stimuli, and two perceptual tasks, a tonal identification task and a lexical task, are adopted to detect lexical effects.

3.1. Stimuli

The experiment has 30 stimuli total, 3 continua (onset pitch, offset pitch, and pitch drop) x 10 variations. Each stimulus consists of a /ti/ syllable, 380 ms in duration, and 80 dB in intensity. These attributes are set as controlled variables based on Yeh's [8] production results and the fact that *f0* is the main cue for tonal perception [9][1]. Pitch ranges, including onset pitch height, offset pitch height, and pitch drop, were synthesized by Praat version 5.1 [1]. In the onset pitch continuum, the pitch ranges of the 10 tonal stimuli are 100-90 Hz, 110-90 Hz, 120-90 Hz, etc. In the offset pitch continuum, the pitch ranges of the 10 stimuli are 160-60 Hz, 160-70 Hz, 160-80 Hz, etc. In the pitch drop continuum, the pitch ranges of the 10 stimuli are 120-60 Hz, 130-70 Hz, 140-80 Hz, etc.

3.2. Participants

30 Hakka speakers (6 males, 24 females; mean age: 39.8 years old) were recruited from the Hsin-Chu area of Taiwan. 30 Taiwan Mandarin speakers (9 males, 21 females; mean age: 26.3 years old) were recruited from Michigan State University.

3.3. Tasks

A tonal identification task and a lexical decision task were designed. In each task, there are 120 trials total (30 stimuli x 4 repetitions). Before the identification task was conducted, a training session of tonal categories was given with minimal pairs, such as fu53 'husband' and fu31 'pants', to Hakka participants because they had never learned these labels in Hakka before, whereas Mandarin speakers had learned the tonal labels, T1, T2, T3, and T4, since elementary school.

In the identification task, the participants were asked to identify if they heard a high-falling tone or a low-falling tone. In the lexical decision task, they were asked to recognize if they heard ti53 'to know' of mï55-ti53 'don't know' or ti31 'emperor' of fong55-ti31 'emperor' in Hakka, or ti51 'younger brother' of ti51ti 'younger brother' or ti21 'bottom' of ti21p^{h}ian51 'film' in Mandarin.

4. RESULTS

The perceptual boundary of each continuum was analyzed to examine the three phonetic factors. Two-sample T-test was then conducted to compare the results of the two tasks because the lexical factor is the only variable between the two tasks.

4.1. Onset pitch continuum

There is a perceptual boundary (50% rate) between high-falling and low-falling tones in both languages, as illustrated in Fig. 1 and Fig. 2. The perceptual boundary is consistently located at stimulus #6, 150-90 Hz. The locations of the perceptual boundaries are slightly different in the two tasks, showing that there are more high-falling responses in the lexical task in both languages.

Figure 1: Hakka results of onset pitch stimuli.

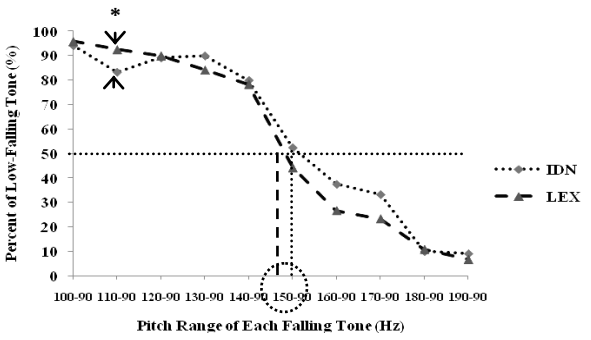
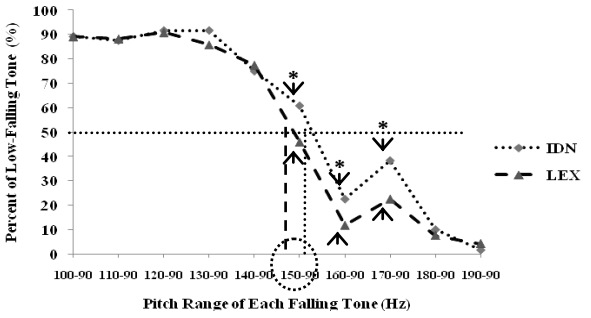


Figure 2: Mandarin results of onset pitch stimuli.



The Two-sample T-Test analysis shows that there is a significant difference in the responses between the identification task and the lexical task at stimulus #2, 110-90 Hz, $t(58)= 2.0095$, $p= 0.0246< 0.05$*, in Hakka; at stimulus #6, 150-90 Hz, $t(58)= 1.7563$, $p= 0.0422< 0.05$*, at stimulus #7, 160-90 Hz, $t(58)= 1.9232$, $p= 0.0297< 0.05$*, and at stimulus #8, 170-90 Hz, $t(58)= 2.1405$, $p= 0.0183< 0.05$*, in Mandarin.

4.2. Offset pitch continuum

No perceptual boundary exists between high-falling and low-falling tones in the offset pitch continuum in either language, as illustrated in Fig. 3 and Fig. 4. Between the two tasks, there are more high-falling responses in the lexical task in both Hakka and Mandarin.

Figure 3: Hakka results of offset pitch stimuli.

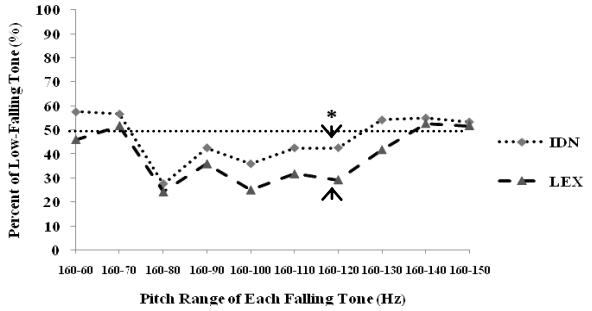
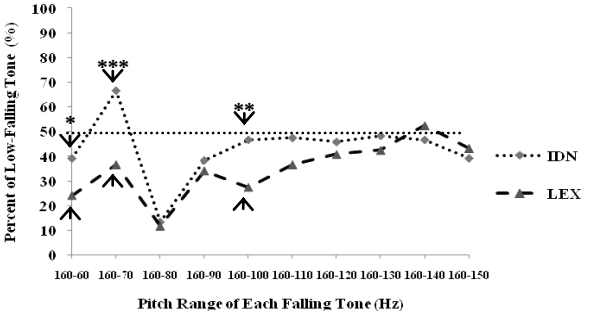


Figure 4: Mandarin results of offset pitch stimuli.



The Two-sample T-Test analysis shows that there is a significant difference in the responses between the identification task and the lexical task at stimulus #7, 160-120 Hz, $t(58)= 1.8230$, $p= 0.0367< 0.05$*, in Hakka; at stimulus #1, 160-60 Hz, $t(58)= 1.9047$, $p= 0.0309< 0.05$*, at stimulus #2, 160-70 Hz, $t(58)= 3.5371$, $p= 0.0004< 0.001$***, and at stimulus #5, 160-100 Hz, $t(58)= 2.4649$, $p= 0.0083< 0.01$**, in Mandarin.

4.3. Pitch drop continuum

There is a perceptual boundary between high-falling and low-falling tones in the pitch drop continuum in both languages, as illustrated in Fig. 5 and Fig. 6. The perceptual boundary is located at stimulus #4, 150-90 Hz, which is consistent with the results in the onset pitch continuum. The locations of the perceptual boundaries are slightly different in the two tasks: there are more high-falling responses in the lexical task in both Hakka and Mandarin.

Figure 5: Hakka results of pitch drop stimuli.

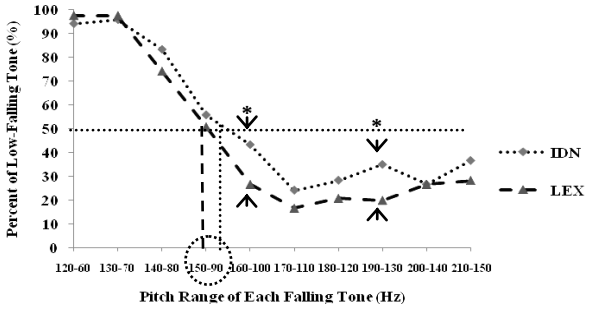
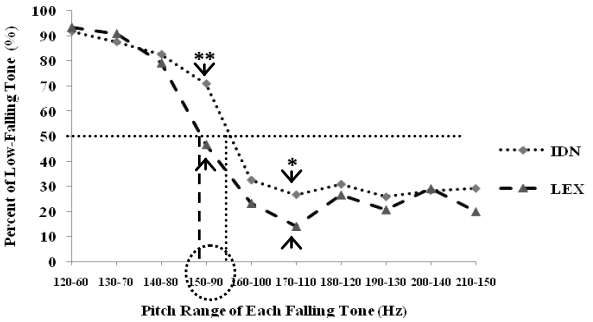


Figure 6: Mandarin results of pitch drop stimuli.



The Two-sample T-Test analysis shows that there is a significant difference in the responses between the identification task and the lexical task at stimulus #5, 160-100 Hz, $t(58)= 2.0097$, $p= 0.0246< 0.05$*, and stimulus #8, 190-130 Hz, $t(58)= 1.9786$, $p= 0.0263< 0.05$*, in Hakka; at stimulus #4, 150-90 Hz, $t(58)= 2.6630$, $p= 0.0050< 0.01$**, and at stimulus #6, 170-110 Hz, $t(58)= 2.0213$, $p= 0.0239< 0.05$*, in Mandarin.

4.4. Results of lexical frequency

The Hakka high-falling word ti53 ‘to know’ is more frequent than the low-falling word ti31 ‘emperor,’ according to Chui and Lai’s [3] corpus study in which the ratio of occurrence frequencies between the high-falling and low-falling words is 114:0. In Mandarin, the high-falling word ti51 ‘younger brother’ is more frequent than the low-

falling word ti21 'bottom,' according to Da's [4] corpus study in which the ratio of occurrence frequencies between the high-falling and low-falling words is 1:0.84.

5. DISCUSSION

There is a perceptual boundary around the tonal variant, 150-90 Hz, in the onset pitch continuum and the pitch drop continuum but no perceptual boundary in the offset pitch continuum. The findings verify hypotheses (1a) and (1c), but reject hypothesis (1b). Onset pitch and pitch drop are crucial to distinguishing high-falling tones from low-falling tones, and pitch perturbations of the two characteristics may cause tonal confusion of the two falling tones. However, offset pitch perturbations exert no significant effects on tonal confusion. The results demonstrate that not all phonetic features are innate perceptual factors for tonal processing.

The significant results between the identification and lexical tasks support hypothesis (1d), showing that the participants categorize the same stimuli in a significantly different manner with different sorts of information. In the lexical task, they tend to choose the frequent words as responses in both Hakka and Mandarin. There are, therefore, more high-falling responses in the lexical task, which conforms to the results of lexical frequency and thus supports the lexical frequency effect. However, the lexical effect exerts a somewhat different influence on the two languages. In Mandarin, the lexical effect is found in eight out of 30 stimuli, and is observed largely among ambiguous stimuli, i.e. those in between the two ends. In Hakka, the lexical effect is found in only four cases, and is more randomly distributed. The differences may be attributed to the additional training session of tonal categories given to the Hakka participants, or the different phonological relations between high-falling and low-falling tones in the two languages.

6. CONCLUSION

This study suggests that phonetic similarity can cause tonal confusion even for adult native speakers, regardless of the phonological status of the falling tones, but not all phonetic features are relevant to tonal distinction and tonal confusion in speech perception. Effects of prosodic contexts on pitch variations are common and substantial, and tonal confusion is supposed to be a common occurrence; however, it seems hardly a problem in daily conversation. Pierrehumbert [7] claims that lexical advantage is the key to speech processing. The Mandarin results show that lexical frequency exerts a significant influence on those ambiguous stimuli. In other words, lexical advantage seems to reduce phonetic ambiguity. This interpretation, however, does not gain strong support from the relatively random pattern of the lexical effect in the Hakka data, probably owing to the methodological or phonological differences. A further investigation is needed to resolve the issue. Nonetheless, our study generally supports the lexical frequency effect on the perception of the two falling tones, although it is still unclear how lexical advantage resolves tonal confusion resulted from phonetic similarity. In sum, this study verifies both the phonetic and lexical effects on tonal perception, supporting Pierrehumbert's exemplar model [7] for speech processing.

7. ACKNOWLEDGEMENTS

We thank two anonymous reviewers for their valuable comments and especially Dr. José Benki for his help in the design of the experiment.

8. REFERENCES

[1] Boersma, P., Weenink, D. 2009. Praat Version 5.1.04. *http://www.fon.hum.uva.nl/praat/download_win.html*

[2] Chao, Y. 1968. *A Grammar of Spoken Chinese*. Berkeley, CA: University of California Press.

[3] Chui, K., Lai, H. 2009. The NCCU corpus of spoken Chinese: Mandarin, Hakka, and Southern Min. *Taiwan Journal of Linguistics* 6(2), 119-144.

[4] Da, J. 2004. Chinese text computing. *http://lingua.mtsu.edu/chinese-computing*

[5] Lo, C. 1990. *Taiwan's Hakka.* Taipei: Taiuan Publishing.

[6] Moore, C., Jongman, A. 1997. Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America* 102(3), 1864-1877.

[7] Pierrehumbert, J. 2003. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46, 115-154.

[8] Yeh, C. 2009. Acoustic correlates of Tone 3 and Tone 4 in Mandarin. *Journal of the Acoustical Society of America* 125(4), 2751.

[9] Yip, M. 2002. *Tone*. Cambridge University Press.

[1] Both reviewers point out that high-falling tone is short in duration. The longer duration setup in the experiment, however, did not disfavor or lower the accuracy percentage of high-falling responses.