THE SIGNIFICANCE OF 'SECONDARY CUES' FOR TONAL IDENTIFICATION IN FUZHOU

Cathryn Donohue

Australian National University, Australia
cathryn.donohue@anu.edu.au

ABSTRACT

This paper examines the role that a change in phonation and a slight contour in fundamental frequency (f0) have for the perception of the three phonologically level tones in the Fuzhou variety of Chinese. An experiment was devised whereby synthetic tokens with modified f0 heights were included among natural tokens and presented to native listeners who were asked to identify specific words from a given set when they heard them. The results of the experiment show that both phonation and slight f0 rise/fall are significant factors for the correct identification of phonologically 'level' tones in Fuzhou.

Keywords: tone perception, tone identification, Fuzhou, phonation, f0 contour

1. INTRODUCTION

Chinese languages are mostly celebrated for their complex tonal contours; however, the use of non-modal (especially breathy) phonation is well attested in the Wu dialects, e.g. [8], and can occur in other varieties as well. Even though non-modal phonation typically occurs with a specific tone or tones independent of segmentals, it is usually considered no more than a secondary cue as the tones are otherwise distinct, distinguished by their pitch contours. In fact, the perceived importance of the non-modal phonation is such that it is often not even mentioned in auditory descriptions and thus subsequently omitted from phonological studies.

In this study, we present results from a tonal identification experiment investigating the role of f0 contour and phonation in the perception of phonologically level tones in the Fuzhou variety of Chinese.

1.1. Fuzhou

Fuzhou is a Min dialect spoken mainly in and around Fuzhou city, north-eastern Fujian Province, China. It has seven citation tones and a complex right dominant tone sandhi system, though it is perhaps most famous for its vowels that 'alternate' between citation and sandhi contexts. There are a number of descriptive works dating back to 1930 [9] which have informed the main phonological works that address the tonal complexities of Fuzhou, e.g. [1, 2, 7, 10].

While the available descriptions differ significantly on the exact shape/height of the pitch contour, the theoretical works largely agree in the representation of tones 1 and 2 as level tones and tone 3 as a level or even rising/LH tone. An acoustic study of the tones, however, reveals a slight contour to these phonologically level tones [4]. The high level tone 1 shows a clear rise in f0 during the second half of the duration, while tones 2 and 3 have very similar, slightly dipping/falling contours (note that the phonological representation of tone 3 as rising is contrary to its phonetic form). Additionally, Donohue [4] notes a change in phonation for a subset of the tones from modal to a glottalized breathy phonation, a change that was consistent across speakers, but not previously reported in any of the other descriptive studies. Donohue [7, 8] claims that the slight contouring of the f0 and the use of non-modal phonation are perceptual enhancements of the tones and (phonological) registers.

A goal of this work is to examine this claim and assess the extent to which these perceptual enhancements play a role in the identification of Fuzhou's level tones. If tones are conceptualized primarily as pitch contours, then one might expect that the slight contour in f0 or the change in phonation are indeed secondary cues and would not be involved in tonal identification. This is the working hypothesis that will be tested in this study.

2. EXPERIMENTAL PROCEDURE

2.1. Tokens

Recordings were made of a female speaker, aged 30 from Fuzhou city, in a sound-proof phonetics lab at the Australian National University.

Tokens selected for this study were (near) minimal pairs whose f0 contours were as close as possible to the mean normalized values of the speaker (illustrated in figure 1) and as representative of those resulting from a normalization across four speakers carried out in a previous study [4] as possible. The tokens used are shown in table 1.

Figure 1: Mean normalized tonal contours for Fuzhou based on 126 tokens from one speaker. f0 was measured at 0,5,20,40,50,60,80,95 and 100% duration.

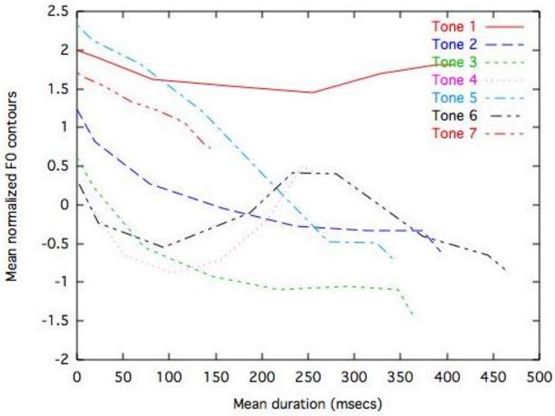


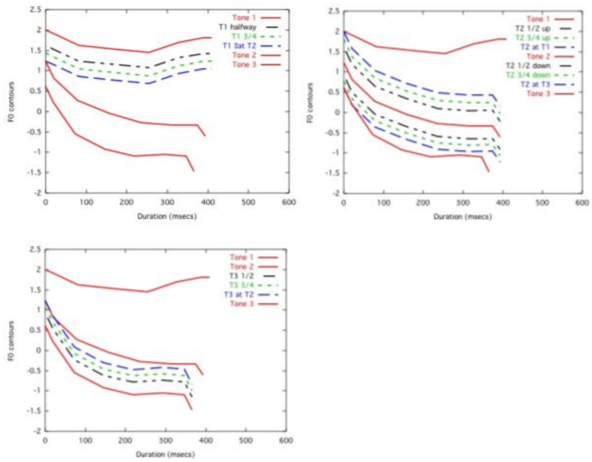
Table 1: Tokens selected for the study.

Tone	Word	Fuzhou	English
1	巴	[pa]	to wish, long for
2	把	[pa]	to handle, grasp
3	霸	[pa]	tyrant, lord
4	百	[paʔ]	hundred
5	爬	[pa]	to crawl, climb
6	第	[ta]	[ordinal prefix]
7	白	[paʔ]	white

2.2. Synthetic tokens

The three ‘level’ tones—tones 1, 2 and 3—were the object of study, so only the tokens produced with these tones were modified. Of particular interest are the pairings of tones 1 and 2 and of tones 2 and 3. The former pair of tones differ in both f0 height, and f0 contour. The latter pair differ in f0 height, but also in that tone 3 is produced with a non-modal phonation; their f0 contours are not significantly different. As the hypothesis being tested is that ‘pitch height’ is the only factor involved in distinguishing these level tones, the natural phonation and f0 contour of the syllables were preserved, only the overall f0 height of the tone was modified, calculated at the point of onset (0% duration). Using the PSOLA function in Praat (http://www.praat.org), the f0 height, calculated from the point of onset of the tone, was altered to create synthetic tokens with the f0 starting at the halfway mark between the points of onset of the tones in a given pair (T1+T2; T2+T3) at three-quarters of the way towards the other tone in the pair and at the point of onset of the other tone in the pair. The modified f0 was calculated from the original measurement intervals (0, 5, 20, 40, 50, 60, 80, 95, 100% duration). The token manipulation resulted in twelve synthetic tokens to be used for the study as shown in figure 2.

Figure 2: Synthetic tokens used in the perceptual study. Solid red contours correspond to (natural) tones 1, 2 and 3. **Top left**: Synthesized versions of tone 1. **Top right**: Synthesized versions of tone 2 both moving up towards tone 1 and down towards tone 3. **Bottom left**: Synthesized versions of tone 3.



2.3. Subjects

The thirteen listeners who participated in this experiment were male and female native speakers of Fuzhou living in the same neighborhood in Hong Kong. They all migrated from the same area near Fuzhou city and had been living in Hong Kong less than 8 years. They spanned three generations, with the oldest listener in his late seventies and the youngest in his late teens. In addition to some knowledge of Mandarin Chinese, as residents of Hong Kong, they also all had some degree of competence in Cantonese. However, all considered Fuzhou to be their mother tongue and were able to communicate in it regularly as the neighborhood in which they live is home to a sizable number of immigrants from the same area of Fujian province.

2.4. Procedure

The experiments were conducted in a quiet room in the home of each listener. The listener was given a sheet with a list of seven words in Fuzhou, corresponding to the core (natural) stimuli presented in table 1, exhaustive of the seven different citation tones (not just the three tones being studied so as not to draw attention to them), as well as an eighth option 'other/none of the above'. The listeners were asked to indicate when they heard one of these words, identifying which word they heard from the list.

The tokens that the listeners heard included the natural tokens bearing all tones 1–7 as well as the synthetic tokens illustrated in figure 2. Additional words (by the same speaker) not from the core stimuli (and segmentally distinct) were also used as filler so that the listeners were not overwhelmed by very similar stimuli back to back. The stimuli were exported into mp3 format and presented using an iPod and a pair of high quality Bose on-ear headphones.

The token identification was noted during the session by pairing identified words with the identifying number of the token illustrated on the iPod. The listeners heard and judged the whole set of stimuli (41 tokens) twice each, identifying the words on the list when and how they judged appropriate. On a few occasions the listener said they weren't sure; these responses were omitted from the data set so as not to count as either positive or negative identification.

3. RESULTS AND DISCUSSION

In order to quantitatively assess the role of phonation and contour, we return to the working hypothesis that was being tested which is that as phonologically level tones, only pitch height should matter for the correct tonal identification. The results of the experiment show that both the slight contour in f0 and the non-modal phonation are significant factors in tonal identification.

3.1. Contour matters

To evaluate the possible role of contour in the perception of tones, the high and mid 'level' tones 1 and 2 were compared. Differing in overall f0 height, these tones also evidenced consistently distinct contours in the acoustic analysis.

There are 52 data points where the synthetic tones for both tones 1 and 2 start at a point exactly half way between the natural tones 1 and 2 (identification of the two synthetic tones, two repetitions, 13 listeners). These synthetic tones have the same f0 height at the point of onset, but distinct tonal contours. We can then set up the hypothesis that the proportion of times, p_0, that a listener would choose the tone with the matching contour (either tone 1 or tone 2) would be 0.5, assuming that contour plays no role in tonal identification.

We look at the difference between the assumed proportion p_0 and the observed proportion p by applying a Z-test for a proportion. The test statistic is given in (1).

(1) $$Z = \frac{|p - p_0| - \frac{1}{2n}}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

In our case, we have n=52, p=45/52=0.865. This yields a Z=5.13. The critical value for $Z_{0.001}$ is 3.29. We can therefore reject, at the 99.9% confidence level, the hypothesis that there is no difference in the proportions (i.e. that contour plays no role in tone identification).

Next we examined the identification of the tones when the f0 is moved three-quarters of the distance towards the other tone (i.e. tone 1 is moved three-quarters of the way down towards tone 2 and tone 2 is moved three-quarters way up to tone 1). In this case, the tones were correctly identified 83% of the time. Thus the f0 height has an effect, but it is a small one. Finally, it is important to see what happens when the tones are moved to the same f0 onset as the other tone. When this is the case, the tones are still correctly classified 81% of the time.

Given that tones are f0 contours, it is not surprising that f0 contour matters. What is surprising is that it is perceptually significant for what are assumed to be phonologically *level* tones. Additionally, as tone 3 has sometimes been represented as LH, the fact that this falling contour is important for tonal identification renders those phonological analyses quite opaque.

3.2. Phonation matters

A similar evaluation was carried out between tones 2 and 3. These tones have similar contours, but differ not only in overall f0 height but also in that tone 3 is produced with a non-modal phonation.

Again, we proceed in a similar manner, comparing tones 2 and 3. We look at the crucial case where the 52 data points with synthetic tones for tones 2 and 3 start at a pitch exactly between tones 2 and 3. These tones primarily differ in their phonation. As in the previous case, we set up the hypothesis that the proportion of times, p_0, that a listener would choose the tone with the matching phonation (either tone 2 or tone 3), would be 0.5, assuming that phonation plays no role in tonal identification. We look at the difference between the assumed proportion p_0 and the observed proportion p by applying the same Z-test for proportion given in (1).

In this case, we have n=52, p=39/52=0.75. This yields a Z=3.47, once again larger than the critical value $Z_{0.001}$ (3.29), allowing us to reject, at the 99.9% confidence level, the hypothesis that there is no difference in the proportions (i.e. that phonation plays no role in tone identification).

At the three-quarter way mark, the tones are identified correctly 62% of the time. When the tones are moved to the onset of the neighboring tone (i.e. tone 2 is moved down to tone 3 and tone 3 is moved up to tone 2), the tones are correctly identified only 56% of the time. However, it is important to note that, unlike for contour, the results should not be symmetrical: tone 2 is more likely to be perceived as tone 3 as its f0 is lowered than tone 3 is to be perceived as tone 2. Indeed, when the tone 2's f0 contour is produced at tone 3's point of onset, it is more often than not perceived as tone 3. However, the same is not true for tone 3 being produced at tone 2's point of onset. This is likely due to the fact that non-modal phonation is produced to varying degrees across the syllable, to the extent that there may be more modal phonation than non-modal phonation across the duration of the word. But while it may be possible to produce tone 3 with little or none of the non-modal phonation, it is never possible to produce tone 2 with it.

These results are particularly important given the omission of non-modal phonation from the descriptions of Fuzhou, with the notable exception of Donohue [4]. Non-modal phonation is not typically phonemic in Chinese languages, which is perhaps why it has been overlooked in previous descriptions of the variety. However, this study has shown that the change in phonation is significant for tonal identification in Fuzhou.

4. CONCLUDING REMARKS

This study has shown that the phonologically 'level' tones in Fuzhou crucially rely on not only the overall f0 height and contour but also on the contrastive use of phonation for their correct identification. While this may seem an obvious point, it is important to recall that for this variety of Chinese at least, with the exception of one acoustic study, non-modal phonation has not been included in the numerous phonetic descriptions of the language, let alone taken into account by phonologists addressing the Fuzhou data.

Phonation has been shown to be important in South East Asian languages, such as Vietnamese and Burmese, but in work on Chinese languages, it has largely been assumed to be a secondary cue and not crucial for phonology or tonal perception. This work shows that for phonologists keen to ground their work in phonetically real or perceptually salient features, both subtle contouring of f0 as well as changes in phonation need to be taken into consideration, rather than relying on simple, impressionistically based features. How to represent the subtle contouring using standard binary phonological features (e.g. H/L, [±raised] etc.) in a language with a large tonal inventory is not clear, but minimally there should be a general statement that in Fuzhou, high tones rise a little and low tones fall a little.

The results of this study suggest that we may benefit from analyzing tone in Chinese as not just f0 contours, but an amalgam of at least f0 and phonation, and possibly other elements not investigated here. While these secondary cues may exist for perceptual enhancement, we have shown that they are nonetheless integral to the accurate identification of the tones.

5. REFERENCES

[1] Chan, L.-L.L. 1998. *Fuzhou Tone Sandhi*. Ph.D. thesis, UCSD.

[2] Chan, M. 1985. *Fuzhou Phonology: A Non-linear Analysis of Tone and Stress*. Ph.D. thesis, U Washington.

[3] Chen, M.Y. 2000. *Tone Sandhi: Patterns across Chinese Dialects*. Cambridge: CUP.

[4] Donohue, C. 1992. *The Phonetics and Phonology of Fuzhou Tones*. Honours thesis, ANU.

[5] Donohue, C. 1999. On the importance of perceptual salience in tonology. *Proceedings of NACCL*. Vol. 1, 1-16.

[6] Donohue, C. 2007. Tones and vowels in Fuzhou. *Proceedings of BLS 33* (to appear).

[7] Jiang-King, P. 1996. *An Optimality Account of Tone-vowel Interaction in Northern Min*. Ph.D. thesis, UBC.

[8] Rose, P. 2000. Independent depressor and register effects in Wu dialect tonology. *JCL* 30(1), 39-81.

[9] Tao, Y.M. 1930. Phonetics of the Foochow dialect. *Minyin Yanjin (BIHP)* 1(4), 445-470.

[10] Yip, M. 1980. *The Tonal Phonology of Chinese*. Ph.D. thesis, MIT.