The role of contour and phonation in Fuzhou tonal identification

Cathryn Donohue
Australian National University

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 the ‘level’ tones in Fuzhou. This paper concludes with a discussion of the implications of the significance of these so-called ‘secondary cues’ for phonology.

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. Rose 2000; Zhu 2006) 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, outside the Wu varieties, the perceived importance of the non-modal phonation is such that it is often excluded from auditory descriptions and thus subsequently omitted from phonological studies.

In this study, we present results from a tonal identification study investigating the role of F0 contour and phonation in the perception of phonologically level tones in the Fuzhou variety of Chinese, showing that they are in fact significant factors in tonal identification.

2. Fuzhou

Fuzhou (Fúzhōu huà (福州話)) is a Mǐn Dōng (閩東) variety of Chinese 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

* This paper is a revised and expanded version of the paper included in the ICPhS 2011 proceedings as Donohue, Cathryn. 2011. The significance of ‘secondary cues’ for tonal identification in Fuzhou. *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, pp. 607–610. Many thanks to the audiences of BLS 2009, ASA 2010, the CAP (ANU) seminar series, ICPhS 2011 and PLRT 2011 conferences and the ANU Tone Workshop for important feedback and discussion. Many thanks also to the anonymous reviewers for their very helpful comments and suggestions, and to Pablo Saz Parkinson for help with Gnuplot. All errors and remaining shortcomings remain the sole responsibility of the author.
that ‘alternate’ between citation and sandhi contexts on a subset of the tones. Examples of the vowel alternations are illustrated in table 1.

<table>
<thead>
<tr>
<th>Set A:</th>
<th>Set B:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tones 1, 2, 5, 7</td>
<td>Tones 3, 4, 6</td>
</tr>
<tr>
<td>i</td>
<td>ei</td>
</tr>
<tr>
<td>ei</td>
<td>ai</td>
</tr>
<tr>
<td>u</td>
<td>ou</td>
</tr>
<tr>
<td>ou</td>
<td>au</td>
</tr>
<tr>
<td>y</td>
<td>oy</td>
</tr>
</tbody>
</table>

Table 1. Alternating vowel pairs in Fuzhou.

The vowels form two groups: a ‘higher’ group (Group A – i, y, u etc.) and a ‘lower’ group (Group B – ei, oy, ou etc.). The Group B vowels are only found with the lower pitch tones, tones 3, 4 and 6, and only in citation/prepausal position. The vowels are all realized as Group A vowels in sandhi position as illustrated in (1)–(3) below (Chao 1934:41).

(1)a. Tone 3 [21]: 氣 [khei 21] ‘air’

(2) a. Tone 4 [23]: 竹 [tøy 23] ‘bamboo’
b. Tone 4 [23] + Tone 4 [23]: 竹節 [ty 5 3ai 23] ‘bamboo section’

(3) a. Tone 6 [231]: 護 [hou 231] ‘protect’
b. Tone 6 [231] + Tone 1 [44]: 護兵 [hu 44 βin 44] ‘guards’

In (1)–(3), the word in the (a) examples not only changes its tone, but also its vowel when it is in sandhi position: the first syllable in the disyllabic expression in the (b) examples.

There are a number of descriptive works dating back to at least 1930 (Tao 1930) which have informed the main phonological works that address the tonal complexities of Fuzhou (e.g. Yip 1980; Chan 1985; Jiang-King 1996; Chan 1998). While the available descriptions differ significantly on the exact shape/height of the pitch contour, the theoretical works largely agree on the representation of tones 1 and 2 as level tones and tone 3 as a level or even rising/LH tone (e.g. Chan 1985). An acoustic study of the tones, however, reveals a slight contour to these phonologically level tones (Donohue 1992).

Figure 1 shows that the high level tone 1 has 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 (1992) 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. Donohue (1999, 2007) claims that the slight contouring of the F0 and the use of non-modal phonation are perceptual enhancements of the tones and (phonological) registers.
3. Phonological cues

A closer examination of other Chinese varieties reveals similar changes in voice quality and vowel quality with certain tones. Cantonese and Mandarin both show vowel ‘lowering’. For example, in Cantonese the vowel [i] in higher stopped tones is lowered to [e] in the low stopped tone e.g. 謂 [sik] ‘know’ vs. 食 [sek] ‘eat’. Mandarin has a slight diphthongization when a high vowel follows a glide in the low tone, tone 3 e.g. 亀 [kwi 55] ‘turtle’ vs. 亀 [kwēi 213] ‘ghost’ and 侒 [lju 44] ‘escape’ vs. 侒 [ljou 213] ‘willow’. However, the catastrophic shifts of the kind seen in Fuzhou are unique to northern Min dialects within China, and remain a recalcitrant issue for phonologists.

Until Donohue (1992), the use of non-modal phonation with tones 3, 4, and 6 in Fuzhou was not discussed. Some might dismiss the phonation change as an epiphenomenon of the lower F0 range. Indeed, the use of creaky voice with low tone syllables in Mandarin and Cantonese (e.g. Belotel Grenié & Grenié 1994, 1997) is claimed to be “redundant”, not distinctive; improving the speed of perception, but not the identification of the tone.

In an earlier study, Khouw and Ciocca (2007) conducted an investigation into the acoustic and perceptually salient aspects of tone in Cantonese. They noted that vowel and voice quality were potential candidates cueing tonal perception that may require investigation. Taking the contour tones into account, they showed that direction and magnitude of F0 change were relevant for tonal production and perception, but level tones were separated by average F0.

With the exception of Donohue 1992, the fact that phonation and the slight contouring of F0 in the level tones in Fuzhou is largely ignored in both phonetic and phonological work.
suggests that it is a fair assumption that these are either unknown or considered unimportant, secondary cues.

A goal of this work is to examine this claim and assess the extent to which these putative secondary cues may play a role in the identification of Fuzhou’s level tones. If level tones are thought to be distinguished primarily by pitch height, 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.

4. Experimental procedure

Perceptual experiments were devised and run to test this claim. An identification task, rather than a discrimination task, was chosen in order to elicit a categorical response to encourage phonetic coding.

4.1 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 the (near) minimal pairs given in table 2. These were tokens that had been used for an acoustic study of the citation tones in previous work (Donohue 1992) that had selected tokens with unaspirated plosive onsets and no codas. These features made these tokens particularly useful for F0 manipulation with readily identifiable onset and offset points.

<table>
<thead>
<tr>
<th>Tone</th>
<th>Word</th>
<th>Fuzhou</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>巴</td>
<td>[pa]</td>
<td>to wish, long for</td>
</tr>
<tr>
<td>2</td>
<td>把</td>
<td>[pa]</td>
<td>to handle, grasp</td>
</tr>
<tr>
<td>3</td>
<td>霸</td>
<td>[pa]</td>
<td>tyrant, lord</td>
</tr>
<tr>
<td>4</td>
<td>百</td>
<td>[pa?]</td>
<td>hundred</td>
</tr>
<tr>
<td>5</td>
<td>爬</td>
<td>[pa]</td>
<td>to crawl, climb</td>
</tr>
<tr>
<td>6</td>
<td>第</td>
<td>[ta]</td>
<td>[ordinal prefix]</td>
</tr>
<tr>
<td>7</td>
<td>白</td>
<td>[pa?]</td>
<td>white</td>
</tr>
</tbody>
</table>

Table 2. Tokens selected for the study.

The exact tokens chosen were those whose F0 contours were as close as possible to the normalized values of the speaker (illustrated in figure 2) and as representative as possible of those resulting from a normalization across four speakers carried out in a previous study (Donohue 1992; figure 1).
4.2 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 was modified 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 and T2; T2 and 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. For example, synthetic tokens for tone 1 were created at 50% of the ‘distance’ (Hz) towards tone 2, three-quarters of the distance towards tone 2 and finally at the point of onset of tone 2. The same was done for the other member of the pair (tone 2’s synthetic tokens) but moving up towards tone 1’s point of onset. The same method was applied to the synthetic tokens in the tone 2/tone 3 pair. The modified F0 was calculated from the original measurement intervals (0, 5, 20, 40, 50, 60, 80, 95, 100% duration) to preserve the original contour. The token manipulation resulted in twelve synthetic tokens to be used for the study as shown in figure 3. In all the plots in figure 3 the solid lines indicate the natural tones 1, 2 and 3. The dash-dot-dot contours are those that start halfway towards the other tone in a given pair, the dotted lines indicate those contours that start 75% towards the
onset of the neighbouring tone and the dashed line shows those contours that have been shifted 100% towards the neighbouring tone and now start at the point of onset of the other tone in a given pair. Thus the first plot shows the tone 1 tokens being shifted down towards tone 2. The second plot contains all the manipulated tone 2 tokens, showing those contours that have moved up towards tone 1 as well as those that have moved down towards tone 3. The final plot shows the tone 3 tokens that have been shifted up towards tone 2.

![Graphs](image)

**Figure 3.** Synthetic tokens used in the perceptual study. Solid contours correspond to (natural) tones 1, 2 and 3. Dash-dot-dot lines represent the halfway mark between onsets in a given pair; dotted lines are 75% towards the onset of the ‘other’ tone and dashed lines start at the onset of the other tone in the pair. **Top left:** Synthesized versions of tone 1 moving towards tone 2. **Top right:** Synthesized tokens of tone 2 both moving up towards tone 1 and down towards tone 3. **Bottom left:** Synthesized tokens of tone 3.

### 4.3 Subjects
The thirteen listeners who participated in this experiment were male and female native speakers of Fuzhou living in the same neighbourhood 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 neighbourhood in which they live is home to a sizable number of immigrants from the same area of Fujian province.

4.4 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 Chinese, corresponding to the core (natural) stimuli presented in table 2, illustrating all seven citation tones (not just the three tones being studied so as not to draw attention to them), as well as an eighth option ‘other/note 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 began with seven natural tokens bearing tones 1–7 to illustrate the speaker’s range and full tonal paradigm to the listener. Following there were synthetic tokens interspersed with natural tokens, including tokens with tones 4–7 so that the listeners were not overwhelmed by very similar stimuli back to back, as well as to reinforce the speaker’s other tones range and to try to draw the listener’s attention away from just tones 1–3. The stimuli were presented using an iPod and a pair of high quality Bose on-ear headphones.

The participants were given the following instructions: You will hear a series of words in Fuzhou. Please identify the word that you hear from the characters in the set below.] The set included characters illustrating all seven tones as identified in table 1, and additionally the ‘Not sure/don’t know’ category.

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 (the eighth response); these responses were omitted from the data set so as not to count as either positive or negative identification.

All procedures were performed in compliance with relevant laws and institutional guidelines and with the approval of the appropriate institutional committee.

Table 3 presents the data obtained in our perceptual study. The rows correspond to the various tokens presented to the subjects, including both real and synthetic tones. The natural tokens on which the synthetic tokens were based are included on the left. The columns represent the tonal categories the subjects identified the tones with (including “unsure”, when subjects could not decide).

---

1 There were only eight Not sure/don’t know responses. These occurred when Tone 1 was produced at Tone 2’s onset (2), when Tone 2 was produced halfway towards Tone 1 (2), when Tone 2 was produced at Tone 3 (2) and when Tone 3 was produced three-quarters of the way towards tone 2 (2). The confusion with the latter two pairs reinforces the role phonation plays in tonal identification; the confusion with the former two suggests that average pitch may be an important feature in identifying Fuzhou tones that warrants further investigation.
How the token was perceived:

<table>
<thead>
<tr>
<th>Tone</th>
<th>Tone 1</th>
<th>Tone 2</th>
<th>Tone 3</th>
<th>Unsure</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 ½ Down</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>T1 ⅔ Down</td>
<td>25</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>T1 at T2</td>
<td>24</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>T2 ⅔ Up</td>
<td>21</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>T2 ⅔ Up</td>
<td>3</td>
<td>19</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>T2 ⅔ Down</td>
<td>4</td>
<td>20</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>T2 at T3</td>
<td>0</td>
<td>26</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>T3 at T2</td>
<td>0</td>
<td>8</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>T3 ⅔ Up</td>
<td>0</td>
<td>9</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>T3 ⅔ Up</td>
<td>0</td>
<td>4</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>T3 at T2</td>
<td>0</td>
<td>0</td>
<td>26</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3. Data obtained in the perceptual study. The tones identified in the first column are the natural tones used as the basis for the synthetic tones in that row.

4.5 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 pitch and the non-modal phonation are significant factors in tonal identification. The results of the raw data are given in figure 4.²

Figure 4 represents the tokens that were played to the listeners during the identification task and how the listeners identified them. Tones 1, 2 and 3 each have real/perceived columns and are located on the x-axis. The y-axis is number of responses for that category; the size of the bars in the histograms reflects the number of tokens/responses, split according to their various categories (natural and synthetic tones, as well as the ‘unsure’ response). Note that the placement of the ‘real’ and ‘perceived’ stacked histograms side by side does not imply a one-to-one correspondence between individual elements (input tokens and responses) in the various categories. For example, the black stripe in the ‘Tone 3: perceived’ column reflects the four “unsure” responses obtained, half of which corresponded to the ‘T3@T2’ token and half to the ‘T3_1/2UP’ token.

² Many thanks to an anonymous reviewer for suggesting the inclusion of this stacked histogram in addition to figure 5.
Figure 4. A graphic representation of the data sorted by tones shows the different tokens used in the task (‘Real’), as well as the responses for that group of tokens (‘Perceived’) in the immediately adjacent column for tones 1, 2 and 3.

Figure 5 is another graphic representation of the results. It takes the three variables – Tone (F0), Contour, and Phonation – and compares the real/actual values with the perceived values and in the final column shows the difference between the two. For example, the ‘real tone’ refers to the actual F0 of the token, which is why it is divided into so many parts, as it covers natural and synthetic tokens. The perceived tone refers to how the speakers identified the tokens they heard (if they classified them as one of the three level tones) and the final plot on the top row shows the difference between the real and perceived ‘tone’. In this case, it is not especially illuminating as there has to be a difference between the actual F0 and the perceived tone given that nature of the doctored F0 and the categorical perception of pitch. The second and third rows present similar results for contour and phonation. The values for contour are either ‘1’ for a tone 1 like contour or ‘2’ for a tone 2/3 contour. All the tokens in this study have either a 1 or 2 value for contour, and are perceived as tone 1, 2 or 3, so would receive a 1, 2, 2 respectively for the perceived contour value. The difference between the actual contour and the perceived contour is nearly zero, as shown by the tall bar above ‘0.0’ in the final plot in the middle row, and the short bar above ‘1.0’, meaning that for most responses, there was no difference between the real and perceived contour. Finally, the third row presents the results for the phonation. Like contour, all tokens either have (‘1’) non-
modal phonaion or not (‘0’), and are perceived as having non-modal phonaion if identified as tone 3 (‘1’) or not if identified as tone 1 or 2 (‘0’). The difference between the actual and perceived phonaion is comparable to that of contour, with slightly more mistakes as indicated by a taller bar above the ‘1.0’.

Figure 5. A graphic representation of the data sorted by features shows the difference (final column) between actual and perceived ‘tone’ (first row), actual and perceived contour (second row) and actual and perceived phonation (third row) and clearly indicates that contour and phonation were nearly always correctly identified by the listeners.

4.6 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).

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

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 typically thought of as F0 contours, it is not surprising that the 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. Future studies, however, might benefit from additionally synthesizing the tokens at intervals not determined by relative F0 onset but rather the average F0 for the whole tone, despite the possible flip-flopping of the onsets, to see if the contour distinguishes the tones given comparable average pitches. This assumes that average pitch height may be an important factor for tonal identification in Fuzhou, as has been shown to be the case for other varieties of Chinese (e.g. Khouw & Ciocca 2007).

### 4.7 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 (75% of the distance towards the neighbouring tone), the tones are identified correctly 62% of the time. When the tones are moved to the onset of the neighbouring 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 throughout 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 majority of descriptions of Fuzhou. 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.3

5. 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.

3 We could have chosen to evaluate the null hypothesis that contour and phonation play no role in tone identification by means of a standard chi-squared test. In the first case, testing for the effect of contour (with N=50), we obtain a chi-squared of 32.8, while in the second test, where we test for the effect of phonation, we obtain a chi-squared value (with N=52) of 14.9. Both far exceed the critical value of 10.83 for the 0.1 per cent level (i.e. \( p<0.001 \)) and one degree of freedom, thus allowing us, once again, to firmly (at > 99.9% confidence) reject the null hypothesis.
The results of this study suggest that we may benefit from analyzing tone in Chinese as not just F0 contours, but as an amalgam of at least F0 and phonation, and possibly other elements not investigated here that can directly interact with segmentals. While these secondary cues may exist for perceptual enhancement, we have shown that they are nonetheless integral to the accurate identification of the tones.

References


Donohue, Cathryn. 2007. Tones and vowels in Fuzhou. To appear in *BLS 33*.


