

UCLA

Working Papers in Phonetics

Title

WPP, No.110: Registers in tonal contrasts

Permalink

<https://escholarship.org/uc/item/12w9g1h8>

Author

Kuang, Jianjing

Publication Date

2012-09-21

Registers in tonal contrasts¹

Jianjing Kuang

kuangjianjing@ucla.edu

Abstract

This study revisits the issue of tonal registers by exploring the cues used in producing and perceiving the five level tones of Black Miao. Both production and perception experiments show that non-modal phonations are very important cues for tonal contrasts. Two different kinds of non-modal phonations that either enhance pitch contrasts or provide an additional contrastive cue divide tonal levels into several registers. Benefiting from more than one cue, 11, 33 and 55 are well distinguished in the tonal space; by contrast, 22 and 44, only contrasting in pitch, are the most confusable tones. The tonal registers model can explain the different uses of non-modal phonations across languages.

1. Introduction

One of the central issues in the representation of tones is how many contrasting levels are needed. For African languages, usually a two-level representation (H vs. L) is enough; and a three-level (H, M and L) representation can cover many other tonal languages. However, three levels have been found not sufficient to account for some languages. Therefore, tonal theory (Yip 2002) has incorporated tonal register features such as [+/- upper], which divides the entire pitch range into two sub-ranges, which can each host two pitch levels: H vs. L. Thus four contrasting levels are possible in a binary feature system. For example, the four Cantonese level tones can be represented in this way:

11 [-upper, L]	33 [+upper, L]
22 [-upper, H]	55 [+upper, H]

This theory is insightful in providing underlying forms for the surface tone representations. Such registers are historically associated with voicing contrasts of onsets. Tones belonging to the [-upper] register originated from voiced onsets, whereas [+upper] register tones are from voiceless onsets.

¹ This work was presented in The Third International Symposium on Tonal Aspects of Languages (TAL), Nanjing, 2012

However, this kind of tonal register theory has to face two major challenges. First, when languages have a fifth level tone, the representation would have to allow an M tone. Then the system becomes a six-level representation, which is typologically not attested.

More importantly, tonal features are supposed to be the knowledge of native speakers/listeners, so that people should be able to use this feature to produce and perceive the tonal contrasts of their languages. For example, according to the tonal theory, Cantonese listeners should be able to perceive the 22 vs. 33 contrast very well, as they are distinguished by upper vs. lower registers and also by H vs. L. However, a perception experiment on the confusability of Cantonese tones (Mok and Wong 2010) showed that tones 22 and 33, which are the most similar in surface acoustic signals, are also the most confusable in perception. In their language survey, (Mok and Wong 2010) observed that young speakers frequently merge these two tones. A first language acquisition study also showed that children did not acquire this contrast until ten years old (Ciocca and Lui 2003). So the [+/-upper] tonal register does not provide a perceivable cue for tonal recognition and acquisition. We thus ask, if speakers are not relying on underlying registers to contrast tones, how do they succeed in contrasting multi-level tones?

To maintain a multi-level pitch contrast is extremely hard for people because of the limitations of production and perception. On one hand, the pitch range used in normal speech is fairly small, usually less than 100 Hz (Baken and Orlikoff 2000); on the other hand, the JND for tone is not less than 9 Hz (Silverman 2003), and languages usually require a much larger difference than the JND to maintain a phonological contrast. For example, tones 22 and 33 in Cantonese are very confusable, but their F0 difference is still about 20-30 Hz (Mok and Wong 2010). Therefore, even a three-level contrast is very hard to maintain in a 100 Hz range, not to mention a fourth or fifth level.

Furthermore, there is no evidence showing that tonal language speakers are more talented in pitch discrimination. For example, English listeners perform equally well with Mandarin listeners in non-speech pitch discrimination (Bent et al. 2006); French listeners even perform slightly better than Trique listeners in a pitch discrimination task (DiCanio 2008). The latter case is very striking, as Trique is a four-level tone language. These studies suggest that tonal perception is a quite different process from pitch perception.

Therefore, the question asked in this paper is, given normal hearing and speaking ability, how can native speakers produce and hear multiple contrasting level tones? I will try to answer this question by exploring the tonal production and perception of a language with five-level-tone contrasts, the most contrasting levels to our knowledge (Edmondson and Gregerson 1992). I will argue that tonal contrasts can be so much more than pitch contrasts. When pitch contrasts get

crowded, other cues will be involved to enhance the contrasts (e.g. pitch contour, phonation cues).

The five-level-tone language that will be discussed in this paper is a Black Miao dialect, called Qingjiang Miao (Ch'ing Chiang Miao). This dialect is spoken at Shidong Kou (Shih-Tung-K'ou), Taijiang (T'ai-Kung) county of Guizhou (Kweichow) province. This dialect was first documented by Fang-Kuei Li in the 1940s, and since then has been the most famous five-level-tone language in tonal studies (e.g. Edmondson and Gregerson 1992, Yip 2002, among many others). I went to the same village to conduct the experiments reported here.

According to Li's transcription, there are eight tones (I-VIII) in this dialect; five of them are level tones, two rising and one falling (using Chao's tonal representation), as shown in Table 1:

Table 1. *Black Miao tonal system.*

I	II	III	IV	V	VI	VII	VIII
44	51	55	22	45	33	13	11

Referring to the historical origins, these eight tones can be divided into two registers (Edmondson and Gregerson 1992, Chang 1947):

upper: **44**, **55**, 45, 13

lower: 51, **22**, **33**, **11**

We can see that two levels are in the upper range and three levels are in the lower range. If [+/-upper] register helps, we would expect 44, 55 should be more distinguishable from 22, 33 and 11. However, [+/-upper] register is not helping in any way for the tones with the most similar pitch values: e.g. 44 vs. 55, 22 vs. 33. Therefore, our study consists of two parts: First, perceptibility tests for the eight tones were conducted to confirm that native listeners are able to hear the contrasts. Second, a production experiment was conducted to examine whether the same native speakers are able to produce the contrasts, and how.

2. Perceptibility of tonal contrasts

The perceptibility tests comprised two perception tasks: identification and discrimination. The goal of these experiments is to determine whether native listeners are able to hear the tonal contrasts in Black Miao, and to examine how these tones are distributed in a perceptual space.

2.1 Methods

2.1.1. Stimuli

The stimuli were a minimal set of eight real monosyllabic words with [pa]. Since there is no database source for an accurate estimation of word frequency for this language, I controlled the syllable type frequency by choosing the minimal set for [pa], which was the most easily recognized and produced by the speakers. All of the test words are frequently used in native speakers' daily life: /pa44/ "send", /pa51/ "drop", /pa55/ "(water) full", /pa22/ "net", /pa45/ "pig", /pa33/ "fail", /pa13/ "father", and /pa11/ "drive away (duck)". I also designed a procedure to overcome any potential lexical frequency bias.

A male native speaker produced a single repetition of each of these words in isolation. This male speaker had a good education background and used to be a Black Miao language teacher. In his productions, he has a good distinction among all of his eight lexical tones. He also recorded the experimental instructions in Black Miao. In the instructions, these monosyllabic targets were explained in Black Miao and used in appropriate contexts so that the subjects would unambiguously understand these words. For example, they would hear "/pa51/, as in 'I dropped my money'" (in Black Miao). The F0 tracks of the stimuli are shown in Figure 1.

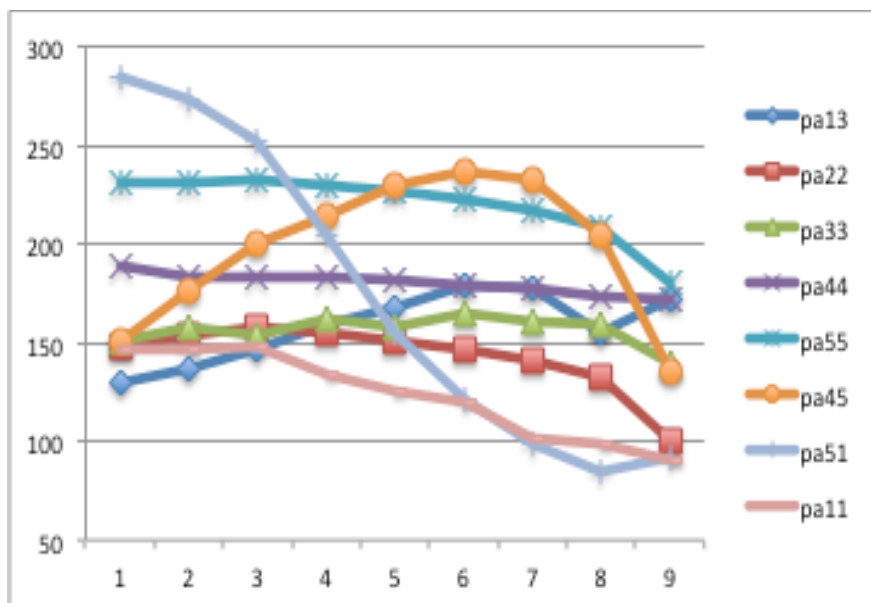


Figure 1: F0 values of the stimuli.

2.1.2. Procedures

The perception experiment was run by a Matlab script on a laptop. The experiment had three phases. The first phase was a familiarity phase, during which subjects were asked to listen to the audio introductions in order to familiarize the test words. They were told that they would hear one of these eight words in each trial. This was to force listeners to pay attention to the phonetic details of these words and to overcome any prior bias about the test words. The experiment thus created the same expectation for all the words. The instructions could be heard as many times as needed until a listener fully understood and memorized the words that would be presented in the following test. When they were ready, they were asked to produce the eight words by themselves first, and each word was repeated twice. This was to make sure these words were fully accessible for them.

The second phase was an identification task. In each trial, an audio introduction "please listen to me carefully" (in Black Miao) by the speaker was played. This was to provide a reference pitch range for the listeners. A single audio target was then presented, and eight test words in both Chinese and Black Miao were displayed on the screen. Subjects were asked to identify which words they had heard by clicking the corresponding button. The task was repeated five times.

The third phase was an AX discrimination task. In each trial, two audio stimuli were presented, and two possible responses, "different" and "same", were displayed on the screen. Subjects were

asked to judge whether the sounds they just heard were same or different words. The stimuli were all possible pairs among the eight stimuli. Thus the "same" trials were 2 instances of a single token. The task was repeated three times.

2.1.3. Subjects

A total of 18 subjects, eight males and ten females, participated in this experiment. Four females, who were not native speakers of this particular Black Miao dialect, were excluded from the current analysis, leaving 14 subjects.

2.2 Results

2.1.4. Identification confusion matrix

Table 2. *Identification matrix for all listeners.*

Target	Perceived							
	T11	T13	T22	T33	T44	T45	T51	T55
T11	85%	3%	0%	2%	0%	2%	7%	2%
T13	0%	88%	0%	7%	0%	2%	3%	0%
T22	0%	7%	19%	5%	64%	0%	3%	2%
T33	0%	10%	3%	81%	0%	2%	2%	2%
T44	0%	2%	7%	3%	76%	0%	0%	12%
T45	0%	5%	2%	0%	0%	93%	0%	0%
T51	3%	2%	0%	5%	0%	0%	90%	0%
T55	2%	0%	2%	3%	7%	0%	2%	84%

Table 2 shows that listeners are able to accurately identify the intended tonal targets in general, except for 22. Contour tones (i.e. 13, 45 and 51) have the best accuracy rates. Among the level tones, the tones at the extreme ends (i.e. 11 and 55) reached better recognition. Interestingly, among the mid-range level tones, 33 has a better accuracy rate than 22 and 44. As 33 is surrounded by 22 and 44 (Figure 1), if it were only distinguished by F0, we would expect it to have the worst identification and there should be a lot of confusion for 22 vs. 33 and 33 vs. 44. However, 33 is rarely confused with either 22 or 44; instead, the tone pair that has a larger F0 difference, 22 and 44, is the most confusable. 22 is mostly perceived as 44 (64%), suggesting that people have difficulties in hearing 22. Could this be because the speaker failed to produce the distinction? Thus we look into individual responses. The 19% accurate responses are mainly from three male and one female listeners. These four listeners show excellent accuracy for all categories (accuracy rate greater than 4/5), indicating that the stimuli do make a distinction between 22 and 44 and that some (though not all) listeners are able to perceive both tones.

2.1.5. Dissimilarity matrix

Compared to the identification task, the discrimination task was much easier for the participants. Indeed, listeners showed generally better accuracy rates than for identification. For example, 70% of responses correctly indicate that T22 and T44 are different, which means people were able to hear a distinction between these two tones. Table 3 is the summary dissimilarity matrix from the discrimination task. Dissimilarity is calculated from the percentage of "different" responses to the tone pairs.

Table 3. *Dissimilarity matrix for all listeners.*

	T11	T13	T22	T33	T44	T45	T51	T55
T11	0.05							
T13	0.94	0.00						
T22	0.93	0.88	0.03					
T33	0.97	0.78	0.95	0.05				
T44	0.98	1.00	0.70	0.98	0.03			
T45	0.94	1.00	1.00	1.00	1.00	0.00		
T51	0.94	1.00	1.00	1.00	1.00	1.00	0.00	
T55	0.95	1.00	1.00	1.00	0.92	0.88	0.88	0.00

Following (Johnson 2003)'s method, we employ Multidimensional-Scaling (MDS) to present this confusion matrix in a low dimensional space. Stress values show that a three-dimensional solution is the best solution, but this is mainly because 51 stands out on the third dimension. As 51 is already quite distinguished in a 2-D solution, shown in Figure 2, we hereby present a 2-D solution instead. As seen in Figure 2, the five level tones are well distinguished in native listeners' minds. Noticeably, the five levels do not contrast along a single dimension. 22 and 33 have similar coordinates on dimension 1 but they are quite distinguished on dimension 2. Other pairs, such as T22 and T11, T44 and T55, contrast on both dimensions. T22 and T44, the two tones that are less contrastive on the first dimension and only little contrastive on the second dimension, are the most confusable tonal pair.

To summarize the perception experiment, native listeners can hear the tonal contrasts very well, except for 22. This result cannot be explained by any pitch-only tonal model, as 33 has much better recognition than 22 and 44, even though the three tones have similar pitch values.

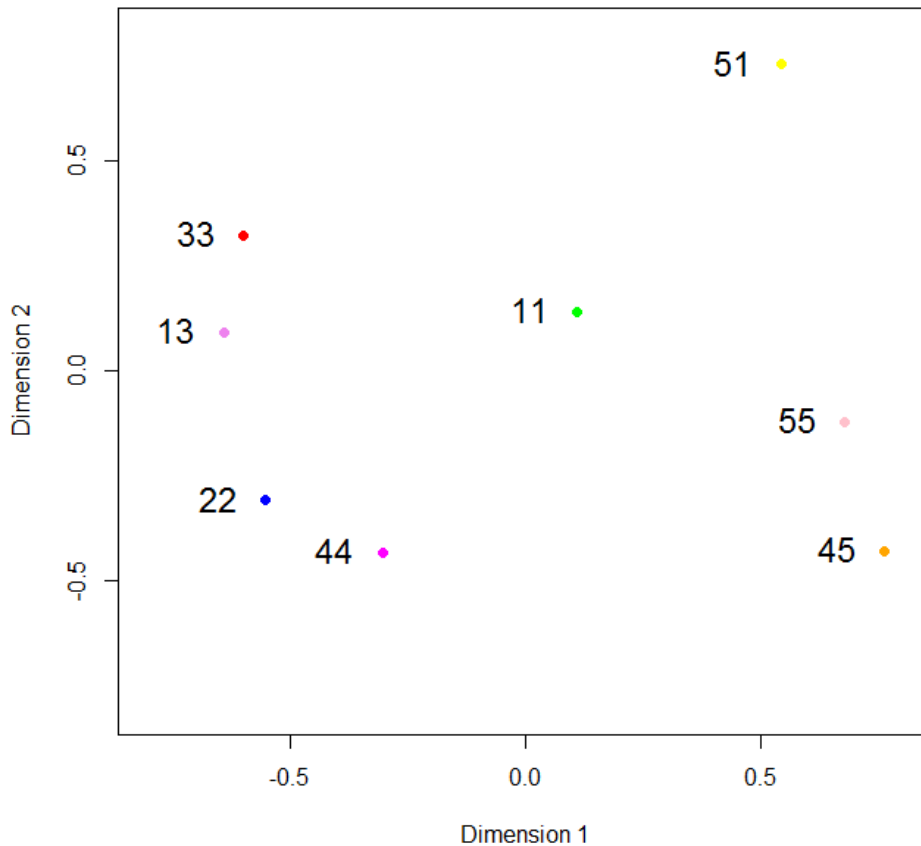


Figure 2: Perceptual space of Black Miao eight tones from discrimination responses.

3. Distinctive cues in producing level tones

From the perception experiment, we learned that the five level tones are largely distinguished in a perceptual space. The corresponding production experiment will further explore how these native speakers achieve these contrasts.

3.1 Methods

3.1.1. Recordings

A wordlist of minimal monosyllabic sets for the eight tones was created based on Li's transcriptions, which were partially reported in (Kwan 1966). These words were first elicited from a fifty-year-old male speaker. 23 minimal or near-minimal sets were confirmed by the speaker. Simultaneous EGG and audio recordings were then collected from 15 native speakers

(ten males and five females). Nine males and five females participated in both the perception and the production experiments. All the speakers were able to understand and speak Southwestern Mandarin. To avoid tone sandhi in continuous speech, the testing monosyllables were read in isolation. Some monosyllables are morphemes that do not normally occur by themselves, but speakers can say them if instructed to.

3.1.2. Measurements

Two sets of measurements were made. The first set includes the traditional F0 related measures (F0, $\Delta F0$, onset and offset) and duration. The second set includes F0 measures plus comprehensive voice measures from both audio and EGG. The idea with this set of measures is to take the entire glottal status into account.

Acoustic measures reflecting different phonation properties were made using VoiceSauce (Shue et al. 2011): H1*-H2* (corrected version by Iseli et al. (2007)), controversially reflecting open quotient of the vocal folds (Holmberg et al. 1995), which has been found to successfully distinguish contrastive phonations across languages (Keating et al. 2011); Amplitude of H1 relative to the amplitudes of the harmonics nearest to F1, F2, and F3 (H1*-A1*, H1*-A2*, H1*-A3*), indicating the strength of higher frequencies in the spectrum, which might be related to closing velocity of the vocal folds (Stevens 1977); Cepstral peak prominence (CPP) (Hillenbrand et al, 1994), reflecting the harmonics-to-noise ratio, which has been found to be an indicator of breathy phonation (Blankenship 2002); H2*-H4*, which is related to vocal fold body tension (Zhang et al. 2011). The EGG analysis in our study is done by EggWorks (Tehrani 2012). Three measures were extracted from the EGG signals: Contact Quotient (CQ), which is defined as the duration of the vocal fold contact during each single vibratory cycle (Rothenberg and Mahshie 1988); Peak Increase in Contact (PIC), defined as the amplitude of the positive peak on the DEGG wave, corresponding to the highest rate of increase of vocal fold contact (Michaud 2004); Speed Quotient (SQ), defined as the ratio between closing duration and opening duration.

3.2 Results

3.1.3. Pitch analysis

A series of pairwise mixed-effect models were used to decide which measures significantly distinguish one tone from another. In the first analysis, average F0, $\Delta F0$, F0 onset, F0 offset and duration were the dependent variables. Average F0 is significantly different between every pair of tones, except for 22 vs. 33. $\Delta F0$ can distinguish contour tones (i.e. 51, 13, 45) from level tones (i.e. 22, 33, 44, 55), and $\Delta F0$ of 11 is also slightly different from 22. F0 onset and offset mostly help distinguish contours with different directions. Duration does not contribute to any tonal

contrasts. Therefore, the only pitch cue for level tones is the average pitch value. Figure 3 shows the average pitch trajectories for eight male speakers.

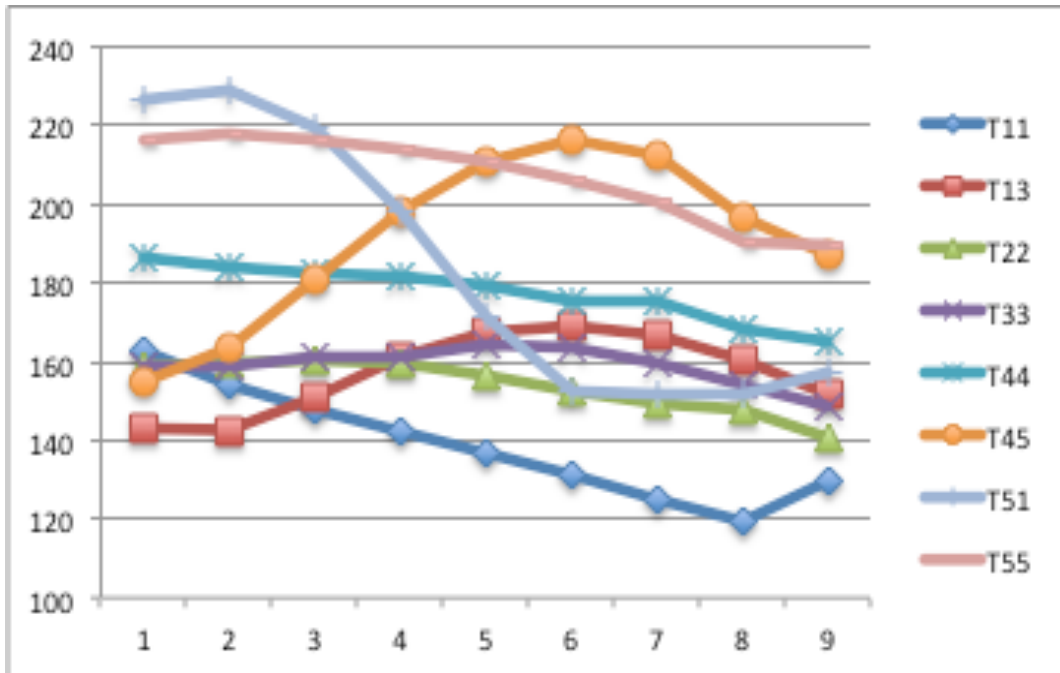


Figure 3: Pitch trajectories for eight male speakers (time normalized).

To see how these level tones are distributed in a physical tonal space, we plot the five level tones by MDS with all the pitch measures. Stress values show that a 2-D solution, shown in Figure 4, is good enough for these data.

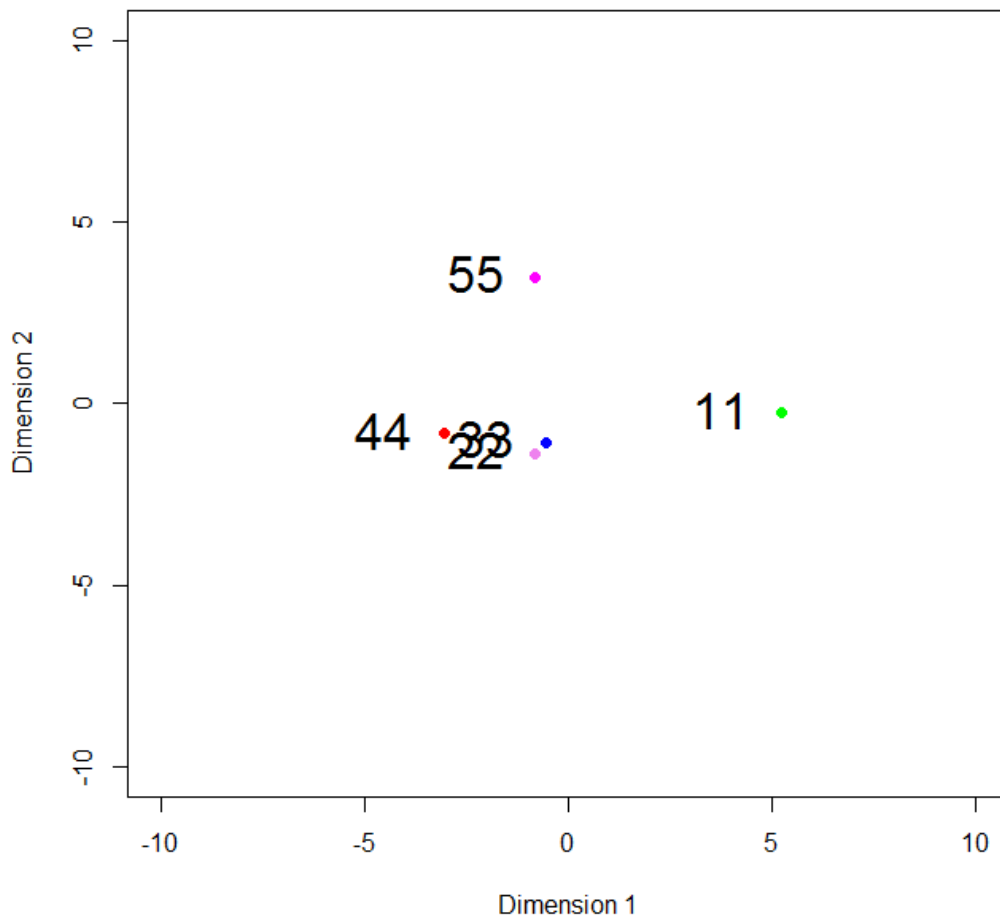


Figure 4: MDS tonal space with pitch measures, level tones only.

Ideally, this production pattern should replicate the pattern shown in the perceptual space (Figure 2), in which the level tones are well dispersed. However, in the pitch-based tonal space (Figure 4), the mid-range level tones collapse together. And 22 and 33 are the most similar pair, which is again not true in perception. Therefore, there must be other important cues that native listeners have relied on.

3.1.4. Phonation cues

In the second analysis, we include all the voice measures in the mixed-effect models. Results show that 33 is much breathier than any other tones, as it has significantly much smaller CQ and greater H1*-H2*, H1*-A1* (Figure 5). On the other hand, 11 and 55 are much creakier than any other tones, as they have greater CQ and smaller H1*-H2*. 22 and 44 have similar voice

quality, which is in between the breathy tone 33 and the creaky tones (i.e. 11 and 55). These results indicate that non-modal phonations are involved in the tonal contrasts.

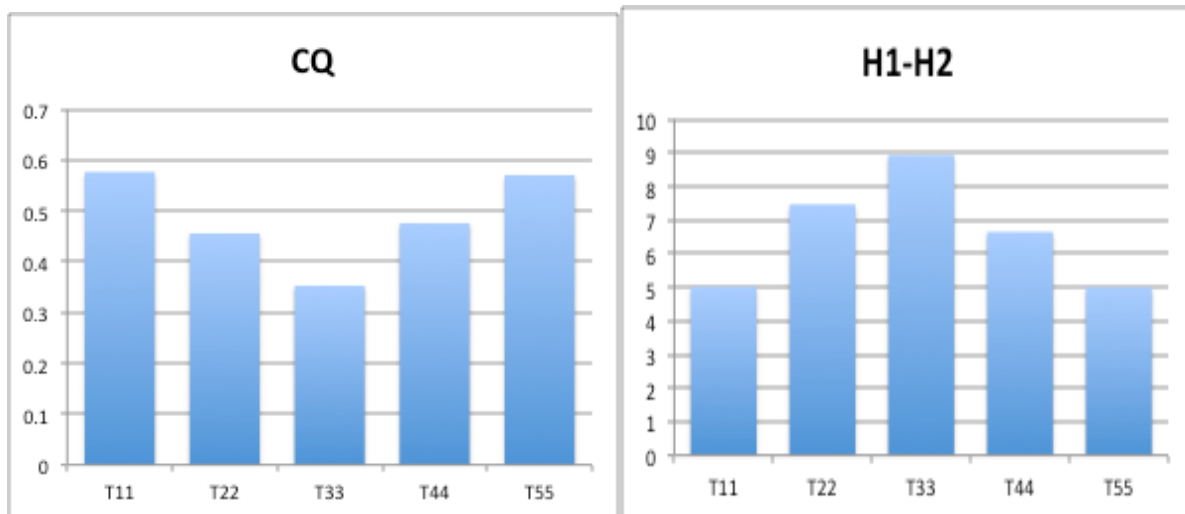


Figure 5: Phonation of the five-level tones.

Incorporating these phonation cues, we regenerate the MDS tonal space (Figure 6). We can see significant improvements from Figure 4: First of all, T33 now is well distinguished from T22 and T44; second, the scale of the space is much bigger than Figure 4, which indicates a better dispersion in general. The new production space now matches better with the perceptual space. This result indicates that non-modal phonations in Black Miao are very important in production, and by inference, also in perception.

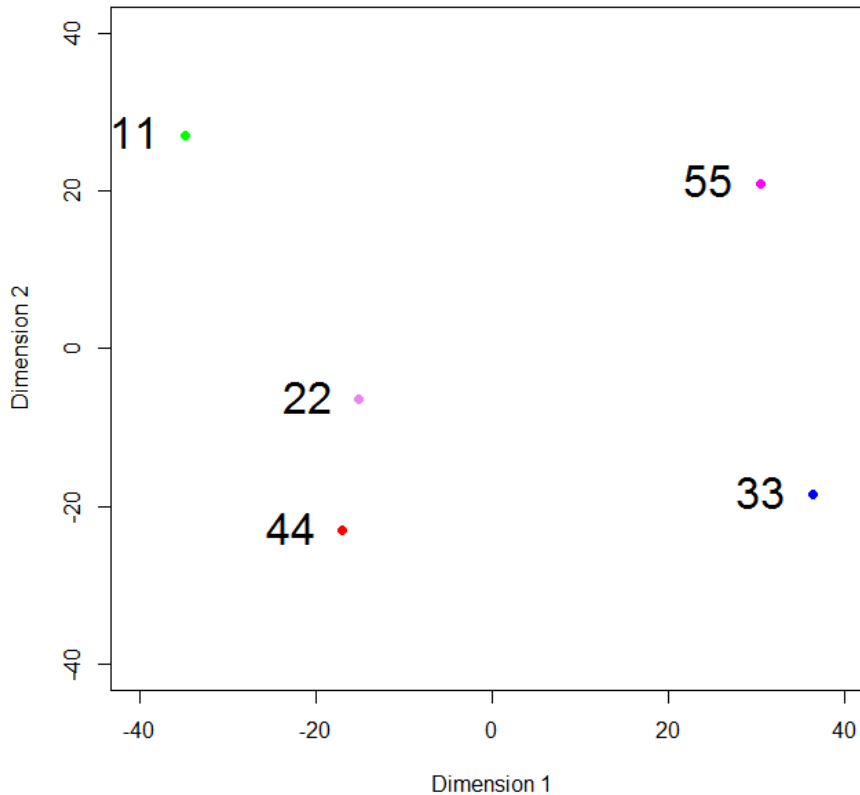


Figure 6: MDS tonal space with pitch and phonation measures, level tones only. Note the scale of Figure 6 is much larger than that of Figure 4.

4. Discussion – tonal register model

In this study, we conducted both production and perception experiments with Black Miao, to explore how native speakers produce and perceive the contrasting five level tones. We confirmed that pitch is not the only cue in tonal contrasts for this language, and non-modal phonations appear to be very important cues in both tonal production and perception. 55 and 11 can benefit from both pitch cues and phonation cues so that they have very good separability. For the mid-range tones that have very similar pitch cues, 33 is distinctive from 22 and 44 primarily by the phonation cue. 22 vs. 44, the tonal contrast with only a pitch difference, is the hardest to produce and perceive. This can be seen in both production and perception maps. Figure 7 generalizes how phonation cues contribute to tonal contrasts in the five-level-tone system.

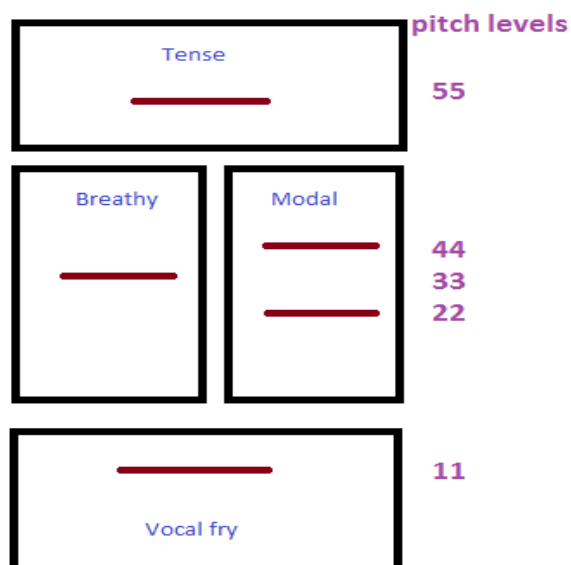


Figure 7: Phonation registers of the five contrasting levels.

In this schema, the five level tones are divided into different registers based on different phonations. The tone with the highest pitch and the lowest pitch form their own registers, and the tones with mid-range pitches can be further divided into two registers: 33 in the breathy register, but 22 and 44 in the modal register. Except for 22 vs. 44, tonal contrasts can benefit from both pitch and phonation cues.

Comparing Figure 6 with Figure 4, the non-modal phonations contribute to the improvement of tonal distinctiveness in two ways: On one hand, the phonation cues enhance the contrasts for 11 and 55, so that the general scale of the production map is enlarged; on other hand, the breathy phonation creates an independent dimension for 33, so that 33 is very distinctive from the other mid-range tones, i.e. 22 and 44.

These two functions reflect the different relationships between pitch and non-modal phonations. The first kinds of non-modal phonations are parts of the pitch scale, such as vocal fry, falsetto and tense. Vocal fry is coarticulated with the lowest pitch range, and falsetto or tense voice is usually associated with the highest pitch range. Referring to Figure 3, the mean F0 of the highest tone is around 220 Hz, which is a remarkably high pitch for male speakers, much higher than the average 175 Hz upper limit of the male speech range (Baken and Orlikoff 2000). The number is important, since when male speakers or singers produce pitches higher than 175 Hz,

they usually have to switch their voice quality into falsetto (Baken and Orlikoff 2000, Keidar et al. 1987). If not, then these high pitches must be produced with tense voice. This results in a greater CQ in EGG signals. Likewise, when pitch goes to the lowest end, e.g. below 75Hz for males, people have to produce these pitches with vocal fry, which also leads to a greater CQ. For more details about physiological differences between vocal fry and other creaky phonations, see Gerratt and Kreiman (2001).

Unlike these pitch-driven non-modal phonations, the second type of non-modal phonations, such as breathy, is independent from pitch. This type of non-modal phonation can create an independent dimension for tonal contrasts, so that tones with similar pitches (33 vs. 22 and 44) but in different registers are rarely confused.

Taken together, there are two types of tonal registers: 1) pitch-driven register; 2) pitch-independent register. Why do languages need these two kinds of registers? This question is related to how many contrasting tonal levels can be achieved in languages. In the spirit of dispersion theory (Lindblom 1986, 1990; Lindblom and Maddieson 1988, Flemming 2004), phonological contrasts should maximize the perceptual differences while minimizing the articulatory effort. Therefore, there are two possible ways to optimize tonal contrast in a five-level-tone language: expand the pitch space for tonal contrasts or add an additional contrastive cue. The pitch-driven phonations can help to produce extreme F0 targets, either super high or low, and thus enhance the perceptual differences for the highest and lowest tones. On the other hand, pitch-independent phonations create an independent dimension for tonal contrasts so that tones with similar pitches can be distinguished from each other. In sum, the well-distinguished five-level-tones of Black Miao can be attributed to both kinds of registers.

The tonal register system proposed in this paper can explain the typologically different use of non-modal phonations across languages. As pitch-driven non-modal phonations are related to realization of extreme pitch targets, they are usually found in low tones or super-high tones. Vocal fry in low tones is very common in languages, the famous cases being Mandarin 213 tone and Cantonese 11 tone (Belotel-Grenié and Grenié 1994, Yu and Lam 2011). Perception experiments (Yu and Lam 2011, Yang 2011) have shown that this non-modal phonation can facilitate tonal recognition for these low tones. Non-modal phonation in super-high tones is less documented, but a few languages that have multiple level tones, such as Yueyang Dialect (Peng and Zhu 2010) and PPhN Thai (Ross 1997), have been reported to have falsetto coarticulated with the highest tones. In all these cases, non-modal phonations are allophonic to tonal contrasts, as they are enhancement cues to pitch. By contrast, pitch-independent non-modal phonations usually are a phonemic dimension in languages. For example, tonal contrasts and phonation contrasts are crossed in Yi (Kuang 2011) and Mazatec (Garellek and Keating 2011). For each

pitch level, there are two (or three) phonation registers. Languages with few contrasting levels can optionally use one or the other kind of registers, but languages with multiple level tones have to take advantages of both kinds.

5. Conclusions

This study revisits the discussion about tonal register by exploring the cues used in producing and perceiving five-level tones of Black Miao. Both production and perception experiments show that non-modal phonations are very important cues for tonal contrasts. Two different kinds of non-modal phonations that either enhance pitch contrasts or provide an additional contrastive cue divide tonal levels into several registers so as to optimize the distinctiveness of the tonal space. Future tonal studies should include the analysis of phonation cues.

Acknowledgements

This work is supported by NSF grant BCS - 0720304 to Prof. Patricia Keating and a summer research award from the UCLA Department of Linguistics. I thank Yan-Liang Shue for VoiceSauce and Henry Tehrani for EggWorks.

References

- Baken, R.J., and Orlikoff, R.F., *Clinical Measurement of Speech and Voice*, Singular Publishing Group, 2000.
- Belotel-Grenié, A, and Grenié, M., "Phonation Types Analysis in Standard Chinese", in *Third International Conference on Spoken Language Processing*, 1994.
- Bent, T., Bradlow, A.R., and Wright, B.A., "The Influence of Linguistic Experience on the Cognitive Processing of Pitch in Speech and Nonspeech Sounds", *Journal of Experimental Psychology: Human Perception and Performance*, 32 (1): 97-103, 2006.
- Blankenship, B., "The time course of nonmodal phonation in vowels", *Journal of Phonetics*, 30 (2): 163-191, 2002.
- Chang, K., "Tones of the Miao and Yao languages" (Chinese), *Shiyusuo Jikan* (Bulletin of the Institute of History and Philology, Academia Sinica) 16: 93-110, 1947

Ciocca, V., and Lui, J., "The Development of the Perception of Cantonese Lexical Tones", *Journal of Multilingual Communication Disorders*, 1 (2): 141-147, 2003.

DiCanio, C.T., *The Phonetics and Phonology of San Martín Itunyoso Trique*, ProQuest, 2008.

Edmondson, J. A. and Gregerson, K. J., "On Five-level Tone Systems", *Language in Context: Essays for Robert E. Longacre*, Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics 107, 555-576, 1992.

Flemming, E., "Contrast and perceptual distinctiveness", In B. Hayes, R. Kirchner and D. Steriade [Ed], *The Phonetic Bases of Markedness*, 232-276, Cambridge University Press, Cambridge, 2004.

Garellek, M. and Keating, P., "The acoustic consequences of phonation and tone interactions in Mazatec", *Journal of the International Phonetic Association*, 41(2): 185-205, 2011.

Gerratt, B. R., and Kreiman, J., "Toward a Taxonomy of Nonmodal Phonation", *Journal of Phonetics*, 29 (4): 365-381, 2001.

Hillenbrand, J.M., Cleveland, R.A. and Erickson, R.L. "Acoustic correlates of breathy vocal quality", *Journal Speech and Hearing Research*, 37: 769-778, 1994.

Holmberg, E.B., Hillman, R. E., Perkell, J. S., Guiod, P.C. and Goldman, S.L. "Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice", *Journal Speech and Hearing Research*, 38:1212-23, 1995.

Iseli, M., Shue, Y.L., and Alwan, A., "Age, Sex, and Vowel Dependencies of Acoustic Measures Related to the Voice Source", *Journal of the Acoustical Society of America*, 121: 2283, 2007.

Johnson, K., *Acoustic and Auditory Phonetics*, 67-77, Wiley-Blackwell, 2003.

Keating, P., Esposito, C., Garellek, M., Khan, S. and Kuang, J, "Phonation contrasts across languages", *Proceedings of ICPhS XVII*, 2011.

Keidar, A., Hurtig, R. R., and Titze, I. R., "The Perceptual Nature of Vocal Register Change." *Journal of Voice*, 1 (3): 223-233, 1987.

Kuang, J., "Production and Perception of the Phonation Contrast in Yi", MA thesis, UCLA, 2011.

Kwan, J.C., "A Phonology of a Black Miao Dialect", University of Washington, 1966.

Lindblom, B., "Phonetic universals in vowel systems", In J.J. Ohala and J.J. Jaeger [Ed], *Experimental Phonology*, Academic Press, Orlando, 1986.

Lindblom, B., "Phonetic content in phonology", *PERILUS* 11, 1990.

Lindblom, B. and Maddieson, I., "Phonetic universals in consonant systems", In L.M. Hyman and C.N. Li [Ed], *Language, speech and mind, Studies in honour of Victoria A. Fromkin*, 62-78, London, Routledge, 1988.

Michaud, A., "A measurement from electroglottography: DECPA, and its application in prosody", *Proceedings of Speech Prosody*, 2004.

Mok, P.P.K., and Wong, P.W.Y., "Perception of the Merging Tones in Hong Kong Cantonese: Preliminary Data on Monosyllables", *Proceedings of Speech Prosody*, 2010.

Peng, J.G. and Zhu, X.N., "Falsetto in Yueyang dialect" (Chinese), *Journal of Contemporary Linguistics*, (001): 24-32, 2010.

Rose, P., "A Seven-tone Dialect in Southern Thai with Super-High: Pakphanang Tonal Acoustics and Physiological Inferences", *Southeast Asian Linguistic Studies in Honour of Vichin Panupong*, 191-208, 1997.

Rothenberg, M. and Mahshie, J. J., "Monitoring vocal fold abduction through vocal fold contact area", *Journal Speech and Hearing Research*, 31:338-51, 1988.

Shue, Y.L., Keating, P., Vicens, C., and Yu, K.M., "VoiceSauce: A program for voice analysis", *Proceedings of ICPhS XVII*, 2011.

Silverman, D., "Pitch Discrimination During Breathy Versus Modal Phonation", In J. Local, R. Ogden, and R. Temple [Ed], *Laboratory Phonology VI*, Cambridge University Press, Cambridge UK, 293-304, 2003.

Stevens, K. N., "Physics of laryngeal behavior and larynx modes", *Phonetica*, 34:264-279, 1977.

Tehrani, H. <http://www.linguistics.ucla.edu/faciliti/sales/software.htm>, accessed on 10 February, 2012.

Yang, R.X., “The Phonation Factor in the Categorical Perception of Mandarin Tones”, Proceedings of ICPhS XVII, 2011.

Yip, M.J.W., *Tone*, Cambridge University Press, 2002.

Yu, K.M., and Lam, H.W., “The Role of Creaky Voice in Cantonese Tonal Perception”, Proceedings of ICPhS XVII, 2011.

Zhang, S.Y., Kreiman, J., and Gerratt.B., “Perceptual sensitivity to changes in vocal fold geometry and stiffness”, *Journal of the Acoustical Society of America* 129: 2529, 2011.