

行政院國家科學委員會專題研究計畫 成果報告

台閩語 non-modal 音質之聲學研究

計畫類別：個別型計畫

計畫編號：NSC93-2411-H-009-020-

執行期間：93年08月01日至95年01月31日

執行單位：國立交通大學外國語文學系

計畫主持人：潘荷仙

計畫參與人員：戴怡欣 黃正宏

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 95 年 1 月 16 日

Voice Quality of Falling Tones in Taiwan Min

Ho-hsien Pan

Department of Foreign Languages and Literatures
National Chiao Tung University, Hsinchu, TAIWAN
hhpan@cc.nctu.edu.tw

Abstract

This study compared the voice quality of high and low falling unchecked and checked tones. Spectral tilt measured as H0-A2 was taken from the syllable midpoint of second, third and four syllables in SVO sentences. Results showed that the voice quality of checked tones was not always creakier than unchecked tones. The voice quality of a syllable at the end of an utterance was not creakier than a syllable at the beginning of an utterance. Narrow focus was a salient factor affecting voice quality. While voice quality of narrow focus syllables became creakier, voice quality of post-focus syllables became more breathy. Focus, a prosodic factor, affected voice quality in a more consistent manner than did lexical tone and syllable position within an utterance in Taiwan Min (Taiwanese).

1. Introduction

1.1. Falling checked versus unchecked tones

Phonologically, checked tones in Taiwan Min can be carried only by CVC syllables of which the final consonant is one of four unreleased voiceless stops, that are bilabial, alveolar, velar, and glottal stops /p, t, k, ʔ/. The presence of final unreleased voiceless stops closes off the vowel nucleus in checked syllables with both an oral and laryngeal closure. Thus, checked syllables in Taiwan Min are known for their short duration. Owing to the short duration of checked syllables, the final tone target (L) in checked high (ML) and falling (HL) tones has been ignored in many previous studies. In these studies, the HL tone is specified as only H while the ML tone is specified as only M.[1-4]. In a fiber optical study, glottalization was observed accompanying the final voiceless unreleased stop, /p, t, k, ʔ/ in Taiwan Min [5].

1.2. Voice quality as a phonological and prosodic distinction

The variation of voice quality during speech can be used to communicate para-linguistic information, such as emotion, attitude, and personality [6, 7]; and linguistic information including phonological contrasts of consonants and vowels [8, 9], and prosodic information [10-13].

Voice quality can be used along with supralaryngeal vocal tract configurations to distinguish between segments and it can also be employed to distinguish between prosodic conditions related to prominence and phrase position [10, 11, 13-15].

This study expanded the phonological contrast employing voice quality from a segmental level to a suprasegmental level. Instead of studying the phonation contrast accompanying vowel and consonant distinctions, this study

investigates the phonation distinction accompanying lexical tonal contrasts at a suprasegmental level.

Instead of using an invasive means, this study investigated voice quality of checked tones through acoustical parameters, including spectral tilt. Spectral tilt, H0-A2, has been demonstrated to be effective in distinguishing a phonological contrast between modal and non-modal voice qualities[8].

It was found that for modal voice, the energy for harmonics in the energy spectrum decrease at a rate of six dB per octave from low frequency to high frequency regions. Thus, energy is high in low frequencies and low in high frequencies. The amplitude difference between first harmonic and the strongest harmonic during second formant measured in H0-A2 is high in modal voicing. As the voice quality becomes breathy, the energy at the high frequency region decreases, and so the difference in energy between low and high frequencies increases. H0-A2 values are high in breathy voice. As the voice quality becomes creakier, the energy for harmonics at the high frequency region begins to increase so much so that high frequency energy is greater than low frequency energy. H0-A2 values in creaky voice are often negative. Generally speaking H0-A2 values are positive and highest for breathy voice, positive and high for modal voice, and lowest and often negative for creaky voice (breathy > modal > creaky).

1.3. Research questions

Two linguistic issues concerning voice quality of checked and unchecked falling tones in Taiwan Min including (1) the voice quality distinction associated with a phonological contrast for unchecked versus checked falling tones, and (2) voice quality related to focus conditions in Taiwan Min, were investigated here in present study.

As far as phonological contrast and voice quality were concerned, since final unreleased voiceless stops for checked tones were accompanied by glottalization, it was hypothesized that voice quality of checked falling tones (H, M) should be creakier and with a flatter spectral tilt than corresponding unchecked tones (HL, ML). In other words, the H0-A2 values for checked tones with creaky voice quality should be smaller than H0-A2 values for unchecked tones.

As far as focus condition and voice quality was concerned, it was hypothesized that the distinction of voice quality for checked and unchecked tones will increase in narrow focus conditions. In other words, the H0-A2 values should be smaller for narrow focused checked syllables (H, M), while the H0-A2 values should be larger for narrow focused unchecked tones (HL, ML).

This study investigated the role that voice quality played in lexical tone and focus distinctions. Instead of using read speech, data elicited through short dialogues was used to ensure the naturalness of speech style.

2. Method

2.1. Speakers

Three male native Taiwan Min speakers, CYS, LYK, and LWS, who were students at the National Chiao Tung University at time of recording, participated in the experiment.

2.2. Corpus

The corpus consisted of five syllables in SVO sentences. The first two syllables formed a subject, while the third syllable was a verb, and the fourth and fifth syllables formed an object. According to a sandhi rule in Taiwan Min, the subject formed a tone sandhi group, while the verb and object formed another sandhi group. Following the tone sandhi rule, the second syllable of the subject carried a juncture tone, while the third and fourth syllable carried a sandhi tone.

Only Taiwanese syllables carrying HL, ML, H, M lexical tones and with initial sonorants followed by non-high vowel and voiceless unreleased final consonants at the position of second, third and fourth syllables were analyzed, as shown in Table 1. The 1264 sentences analyzed were part of a larger corpus with seven juncture tones carried by the second syllable matching with the six sandhi tones carried by the third syllable, and another six sandhi tones carried by the fourth syllable of a sentence.

2.3. Experimental procedure

During the recording, an experimenter was present in a sound treated booth with one of the speakers. To ensure that speakers produced sentences with the intended focus conditions, the experimenter produced sentences such as those listed in Table 2 to elicit sentences with focus on the entire utterance, subject, verb, or object.

Table 1: *Corpus*. σ : syllable

	Subject 2 nd σ	Verb 3 rd σ	Object 4 th σ
High Falling	[ma ʔ]	[liam ʔ] 'pinch'	[a ʔ] 'duck'
Low falling	[lun ʔ]	[ma ʔ] 'scold'	[lua ʔ] 'comb'
High falling checked	[lat ʔ]	[lak ʔ] 'lost'	[lok ʔ] 'bag'
Low falling checked	[lok ʔ]	[lap ʔ] 'paid'	[liap ʔ] 'sore'

To control for focus condition, sentences were produced in one of four focus conditions including broad focus on the whole sentence or narrow focus on subject, verb, or object. Each sentence in four different focus conditions was repeated three times. There were all together 15168 (1264 sentences \times 4 focus conditions \times 3 repetitions) utterances. The order in which the sentences were produced was randomized.

Table 2: *Example of corpus elicitation*: BF: broad focus; NF: narrow focus; boldface: location of narrow focus.

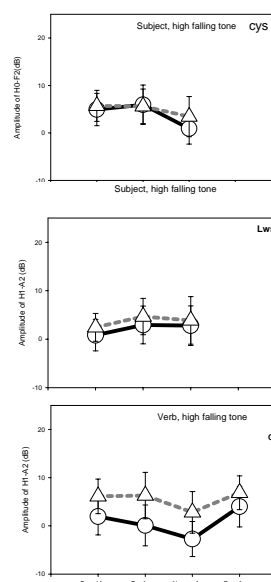
	Experimenter	Speaker
BF on entire utterance	'What happens?'	Grandma hit A-mei.
NF on subject	'Who hit A-mei?'	Grandma hit A-mei.
NF on verb	'What did Grandma do to A-mei?'	Grandma hit A-mei.
NF on object	'Whom did Grandma hit?'	Grandma hit A-mei .

2.4. Data analysis

Only sentences with falling tones on second, third and fourth syllables were analyzed. In total, 15168 sentences were analyzed.

Every sentence was annotated following TW_ToBI conventions using romanization and focus tiers. Next, the onset and offset of the vowel in second, third and fourth syllables were determined from spectrograms, as were the upper and lower bound frequency for the second formant. The duration and the location of the vowel midpoint of second, third and fourth vowels were calculated. Energy spectrums were generated at the midpoint for second, third, and fourth syllables. The amplitudes of H0-A2 and the strongest harmonic during F2, between the upper and lower bound of frequency range for second formant were taken. The spectral tilt, i.e. H0-A2, at the 50% time point was calculated.

3. Results



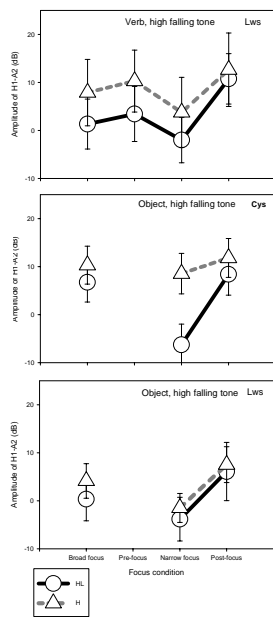


Figure 1: Spectral tilt for high falling tones.

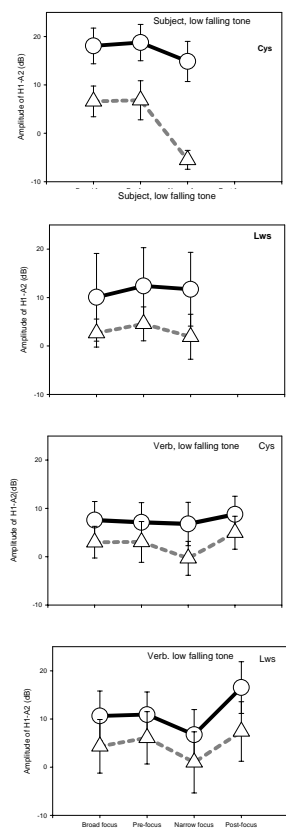


Figure 2: Spectral tilt for low falling tones.

3.1. Effect of Tone

As shown, in Figure 1, when comparing the two high falling tones, the H0-A2 values were higher for the high falling checked tone (H) than for the high falling unchecked tone (HL) at the position of subject, verb, and object. As shown in Figure 2, when comparing the two low falling tones, the H0-A2 values were higher for the low falling unchecked tone (ML) than for the low falling checked tone (M) at the position of subject and verb. However, at object position, the H0-A2 values were higher for the low falling checked tone (M) than for the low falling unchecked tone (ML). In sum, the H0-A2 values for checked tones (H, M) were higher than unchecked tones (HL, ML) at the position of object (H > HL, M > ML). At the position of subject and verb, high falling tones and low falling tones showed different patterns. For the high falling tone, the H0-A2 values were higher in checked tones than in unchecked tones (H > HL), while the reverse pattern was true for the low falling tone where H0-A2 values were higher for unchecked tones than for checked tones (ML > M).

3.2. Effect of focus condition

As shown in Figure 1, among the same high falling tone, the H0-A2 values were the smallest under narrow focus at the position of verb and object, but not at the position of subject. In LWS's production, the H0-A2 values for post-focus syllables were the highest.

As shown in Figure 2, the H0-A2 values among the same low falling tone were the smallest under narrow focus at the position of subject, object and verb. The only exception was the low falling unchecked tone carried by a verb. Again in LWS's production the H0-A2 values were the highest for post-focus syllables carrying low falling tones.

In sum, the H0-A2 values were the smallest under narrow focus with the exception of high falling tones at the position of subject and low falling unchecked tones at the position of verb.

4. Discussion

Results here reveal that the H0-A2 values are both positive and not different for two high falling tones at subject position. The voice qualities of subjects carrying high falling unchecked and checked tones are similar and not creaky like. For subjects carrying a low falling tone, the voice quality in checked tones is creakier than in unchecked tones, since the H0-A2 values for unchecked tones are higher than checked tones. At verb position different patterns were observed for high and low falling tones. For the high falling tone, the voice quality of unchecked tones is creakier than unchecked tones, while the reverse pattern is true for the low falling tone. That is, the voice quality of checked tones is creakier than in unchecked tones. The H0-A2 values for checked tones at the position of object were higher than unchecked tones for both high and low falling tones. Therefore, the voice quality of checked tones is less creaky than unchecked tones at the position of object.

Results here violated the first hypothesis which stated that voice quality of checked tones is creakier like than unchecked tones. The hypothesis is only supported by the voice quality of low falling tones carried by subject and verb. In other words, voice quality of checked tones is not necessarily creakier than unchecked tones.

Among syllables carrying the same tone but under different focus conditions, the H0-A2 values are negative only for syllables under narrow focus. Thus, creaky voice only accompanies narrow focused syllables. Even though H0-A2 values are not negative, they are still the smallest for syllables under narrow focus. Among syllables carrying the same tone, the H0-A2 values for post-focus syllables are the largest. Generally speaking, the voice quality of post-focus syllables is least creaky-like, while the voice quality of narrow focused syllable is either creaky or most creaky-like.

As for the second hypothesis, instead of observing a most creaky voice quality for narrow checked tone and least creaky voice quality for narrow focused unchecked tones, the voice quality for both narrow focused checked and narrow focused unchecked tones becomes creakier regardless of lexical tonal values. Moreover, the voice quality of post-focus syllables becomes breathier to enhance the creaky voice quality for narrow focus syllable. In sum, a prosodic factor, focus, influences voice quality by turning the voice quality of narrow focused syllables creakier, and the voice quality of post-focus syllables breathier.

Prosodic factors influence voice quality in Taiwan Min, just as they influenced voice quality in English. For future studies, the influence of prosodic phrasal position on voice quality in Taiwan Min should be studied to further reveal prosodic influence on suprasegmental features.

5. Acknowledgements

The project is supported by a grant from The National Science Council in Taiwan (93-2411-H-009-020-). I would like to thank Julie McGory for comments on an earlier draft of this paper, Lai-sin Chen and Yi-hsin Tai for assistance in data analysis.

6. References

- [1] Zhang, Z.-X., *Taiwan Minnan Fangyan Jilyue (Notes on Southern Min dialects of Taiwan)*. Taipei: The Liberal Arts Press, 1989,
- [2] Ting, P.H., *Taiwan Yuyan Yuanliou (Sources of Languages in Taiwan)*. Taipei: Student, 1985,
- [3] Cheng, R., "Some notes on tone sandhi in Taiwanese." *Linguistics*, 100: 5-25, 1973.
- [4] Cheng, R., "Tone sandhi in Taiwanese." *Linguistics*, 41: 19-42, 1968.
- [5] Iwata, R., et al., "Laryngeal adjustments of Fukienese stops: Initial plosives and final aplosives", *Annual Bulletin, Research Institute of Logopedics and Phoniatrics*, 13: 61-81, 1979.
- [6] Teshigawara, M. "Voices in Japanese animation: How people perceive voices of good guys and bad guys", in 15th ICPHS, Barcelona, 2003.
- [7] Gobl, C. and A. Ni' Chasaide, "The role of voice quality in communicating emotion, mood and attitude", *Speech Communication*, 40: 189-212, 2003.
- [8] Gordon, M. and P. Ladefoged, "Phonation types: a cross-linguistic overview", *Journal of Phonetics*, 29: 383-406, 2001.
- [9] Blankenship, B., "The timing of nonmodal phonation in vowels", *Journal of Phonetics*, 30: 163-191, 2002.
- [10] Epstein, M.A. "Voice Quality and Prosody in English", in The 15th ICPHS, Barcelona, 2003.
- [11] Epstein, M.A., Voice Quality and Prosody in English, Ph.D. dissertation, University of California, L.A.: L.A., 2002.
- [12] Campbell, N. and P. Mokhtari. "Voice quality: the 4th prosodic dimension", 5th ICPHS, Barcelona, 2004.
- [13] Dilley, L., S. Shattuck-Hufnagel, and M. Ostendorf, "Glottalization of word-initial vowels as a function of prosodic structure" *Journal of Phonetics*, 24: 423-444, 1996.
- [14] Pierrehumbert, J.B. and D. Talkin, *Lenition of /h/ and glottal stop*, in *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, G.Docherty and D.R. Ladd, Editor, Cambridge University Press: Cambridge, p. 90-117, 1992.
- [15] Redi, L. and S. Shattuck-Hufnagel, "Variation in the realization of glottalization in normal speakers" *Journal of Phonetics*, 29: 407-429, 2001.
- [16] Swerts, M. and R. Veldhuis, "The effect of speech melody on voice quality" *Speech Communication*, 33: 297-303, 2001.