

行政院國家科學委員會補助專題研究計畫  成果報告  
 期中進度報告

(計畫名稱)

計畫類別： 個別型計畫  整合型計畫

計畫編號：NSC 98-2410-H-009-040-MY3

執行期間：2009 年 08 月 01 日至 2013 年 07 月 31 日

執行機構及系所：國立交通大學外國語文學系

計畫主持人：潘荷仙

共同主持人：無

計畫參與人員：Professor Pat Keating, 林厚亦，麥家齊、陳彥伶、黃凡、呂紹任、劉佳柔、簡芳誼。

成果報告類型(依經費核定清單規定繳交)： 精簡報告  完整報告

本計畫除繳交成果報告外，另須繳交以下出國心得報告：

赴國外出差或研習心得報告

赴大陸地區出差或研習心得報告

出席國際學術會議心得報告

國際合作研究計畫國外研究報告

處理方式：除列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

中 華 民 國 101 年 05 月 28 日

## 國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以 100 字為限）

實驗失敗

因故實驗中斷

其他原因

說明：

已完成北部漳州、北部泉州、中部漳州、中部泉州、南部混合腔 30 歲以下及 40 歲以上入聲在 citation、雙字詞、句中各母音標記及聲學、EGG 數據分析，目前結果顯示需另外加標入聲尾子音之音段。

2. 研究成果在學術期刊發表或申請專利等情形：

論文： 已發表  未發表之文稿  撰寫中  無

專利： 已獲得  申請中  無

技轉： 已技轉  洽談中  無

其他：（以 100 字為限）

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以500字為限）

#### 學術成就

本研究發現台閩語入聲音節母音帶有 tense 音質，另外在句中，入聲尾之子音具有音長，而非之前假設會消失情況。

#### 技術創新

本研究所使用 VoiceSauce 軟體，原不支援中文系統，在本人多年與 UCLA 共同努力下，於 2011 年 10 月完成中文可使用之完整版本，目前清華大學亦開始使用本軟體。

#### 社會影響

未來台閩語語音處理將會有更會更逼真。

## 國科會補助專題研究計畫項下出席國際學術會議心得報告

日期：\_\_年\_\_月\_\_日

計畫編號	NSC 98- 2410- H- 009 - 040 -MY3		
計畫名稱	台閩語入聲音質之反過濾氣流及聲學研究		
出國人員姓名	潘荷仙	服務機構及職稱	國立交通大學外國語文學系教授
會議時間	100年8月17日至 100年8月21日	會議地點	香港
會議名稱	(中文)國際語音會議 (英文)ICPHS 2011		
發表論文題目	(中文)台閩語入聲電子聲帶圖及聲學研究 (英文) An Acoustical and EGG Study on Checked Tones of Taiwan Southern Min		

## 一、參加會議經過

本次會議為國際語音會議頭一次在亞洲舉辦主辦單位為香港城市大學。本人與台大教授及研究生於會議前一天入住，香港會議中心對面飯店，隨即前往會議地點報到。

第一天 Kohler 教授進行第一場 Plenary Lecture 講述聲音與韻律結構間關係。隨即前往聆聽 Pat Keating 教授介紹 VoiceSauce 軟體，Pat Keating 教授在演講後致謝中特別提到本人擔任該軟體之 Beta User 助其 debug。本會議由於在香港舉辦有多位台灣學者與會，與過去在歐美舉辦，台灣教授出席不出 5 人情況比較大不相同。清華大學謝豐凡教

授一人發表 5 篇文章產量驚人。

第二天 Ishihara 教授講授東京日文焦點與不帶重音之 WH 疑問句間關係，Ishihara 秉持過去對德文再現焦點語氣所做研究功力，再次對日文焦點提出漂亮之處理方式。

第三天 Sara Hawkins 教授發表與不同情境中相同語音線索對語音辨識所產生不同影響。

第四天本人發表台閩語入聲研究成果，會場中座無虛席，甚至有後到者需站立。其中多為華裔學者，演講結束，Edmonson 教授對本人研究結果呼應他的發現甚感興奮。從事福建閩南語泉州方言研究之學者亦提出問題，並要求本人提出假設。

英國 UCL 許毅教授提出以加重與氣候是否有基頻變窄，音長變短現象，對中過各省語言加以分類，會後許教授詢問本人是否可在台灣收集客語資料，與其合作。

本人研究生潘鈺楨發表 Acoustic Cues to Syntactic Ambiguity in Taiwan Mandarin 研究國語歧異據之聲學線索，及呂紹任發表 Taiwan Hakka Languages and TWHK\_ToBI Annotation 研究客語ToBI系統。其中潘鈺楨幾乎無法以英語報告。

## 二、與會心得

由於中國大陸崛起，大量中國學者在會議中出現。這是過去從未看到情形。台灣語音研究應積極上進。學生英文素質應更提升

## 三、考察參觀活動(無是項活動者略)

無

## 四、建議

## 五、攜回資料名稱及內容

Proceedings of ICPHS 2011

# DIALECTAL VARIATION OF VOICE QUALITY IN TAIWAN MIN: AN EGG AND ACOUSTICAL STUDY

*Ho-hsien Pan, Mao-hsu Chen, Shao-Ren Lyu, Yu-chu Ke*

Department of Foreign Languages and Literatures, National Chiao Tung University, TAIWAN  
[hhpan@faculty.nctu.edu.tw](mailto:hhpan@faculty.nctu.edu.tw), [maohsuchen@gmail.com](mailto:maohsuchen@gmail.com), [thousandshine@hotmail.com](mailto:thousandshine@hotmail.com)

## ABSTRACT

This study explored the voice quality of Taiwan Min checked tones 5 and 3 produced by speakers of varying ages from northern, central and southern Taiwan with Zhanzhou, Quanzhou and mixed accents. Comparison between spectral tilt (difference between H1\*-H2\* amplitudes) and closed quotients of the glottal wave form showed that checked tones with creaky / tense voice quality were distinguished from unchecked tones produced by middle aged speakers from central and southern Taiwan but not by young speakers.

**Keywords:** Taiwan Min, sociophonetics, checked tones, sound change, voice quality, EGG.

## 1. INTRODUCTION

In many Sino-Tibetan languages, such as Green Mong, Khmar, Kuai, Jingho, Hani, Yi, and Wa, lexical tones are distinguished by both voice quality and tonal contour [1, 2, 3, 12, 13, 15, 17, 18]. For example, Yi has a tense and lax voice quality contrast as well as f0 tonal contrasts. White Hmong, a Hmong-Mien language, has modal, breathy, and creak phonation contrasts as well as tonal contrasts.

Phonation and tonal contrasts exist in Jalapa languages as well. For example, Mazatec has modal, breathy, and creaky phonation contrasts and tonal contrasts. Also, Indo-European languages, such as Gujarati have phonation contrasts but do not have a tonal contrast.

Esposito [10] by means of Multidimensional Scaling Analysis (MDS) found that listeners' linguistic backgrounds affected the perceptual cues they used to distinguish phonation contrasts. A cross language Electroglottography (EGG) and acoustical study on phonation of Yi, Gujarati, white Hmong, and Mazatec found that spectral tilt, H1\*-H2\*, was effective in distinguishing the phonation types in all four languages [16], whereas EGG data, such as the mean close quotient (CQ\_H), distinguished modal from breathy phonation in Gujarati, tense from lax phonation in Yi, and breathy, creaky, and modal phonation in White Hmong [16]. Languages make use of different acoustical parameters to distinguish phonation types. For example, Mazatec phonations can also be distinguished with H1\*-H2\* and Cepstral Peak Prominence (CPP) as well [11]. In addition to

acoustic and EGG parameters, the time course in which the phonation contrast continues within the vowel also varied across languages. In Yi, the phonation contrast continues through the entire vowel. In Gujarati, the phonation contrast is present in the middle of the vowel, and in Mazatec the phonation contrast exists only in the beginning of the vowel [16].

This study followed this line of study on languages with both phonation and lexical tonal contrasts and investigated the voice quality of two checked tones, 5 and 3, of Taiwan Min. CVC syllables with final unreleased voiceless stops, “checked” tone syllables, are shorter in duration than open syllables, or syllables with nasal codas [7].

A previous fiber optical study on Taiwan Min checked tones 5 and 3 found glottalization accompanying oral closure during articulation of final stops [14]. An inverse filtered oral airflow study found low airflow and a long closed phase in the glottal waveform of checked tones produced by some speakers [19]. As the vocal folds closed off during checked syllables, the airflow decreased and the closed phase of glottal wave forms increased. Acoustical studies using spectral tilt, amplitude of first harmonic minus amplitude of the highest peak of second formant (H1-A2), did not find creaky voice quality for checked tones [20]. It was proposed that H1-A2 may

not be an effective parameter in documenting voice quality of Taiwan Min checked tones. Laryngoscopic study found that all speakers showed glottal fold adduction. Younger speakers showed less or no adduction of ventricular fold or aryepiglottic folds [9]. In sum, fiber optical, laryngoscopic and inverse filtered oral airflow studies all found glottalization for Taiwan Min checked tones, however, the effective acoustical parameters for documenting Taiwan Min checked tones are yet to be determined.

Recent field work studies based on auditory impressionistic data observed ongoing sound change for Taiwan Min checked tone 5. The high onsets of checked tone 5 drop to mid onset as checked tone 3 [4, 6]. To explore how tense / creaky voice qualities were affected by the sound change in  $f_0$  contours, the voice quality of checked tones produced by young and middle aged speakers in Northern Zhangzhou, Northern Quanzhou, Central Zhangzhou, Central Quanzhou, and Southern Mixed dialect regions of Taiwan are documented.

## **2. METHOD**

### **2.1. Subjects**

Following the isoglosses defined by Ang [5], this study recruited native speakers from northern and central Taiwan with Zhangzhou (漳州) and Quanzhou (泉州) accents and

southern Taiwan with a mixed accent. All speakers were able to speak Mandarin. EGG data were recorded from four of 10 speakers, as shown in Table 1.

Table 1. Speakers' backgrounds.

	40-60 years old	20-30 years old
N. Zhangzhou	1Female (1 EGG)	1Male (1 EGG)
N. Quanzhou	2Females (1 EGG)	
C. Zhangzhou	2Females	
C. Quanzhou	1Male	1Female
S. Mixed	1Female	1Male (1 EGG)

## 2.2. Instruments

Glottal Enterprise EGG system EG2-PCX and TEV TM-728II microphones were used to simultaneously pick up acoustic and EGG data. The EGG and acoustic signals were recorded with Audacity onto a laptop and then separated into different channels, inverted and analyzed with EggWorks [8]. The acoustic signals were tagged by Praat to mark vowel boundaries, and then analyzed with VoiceSauce [21].

## 2.3. Corpus

A randomized reading list included 390 monosyllabic tokens that were read in citation

form included the target checked tones, 5 and 3, controlled unchecked tones, 53 and 31 and filler items with rising and level tones.

## 2.4. Procedure

During the recording, two electrodes were placed on four speaker's necks while a microphone was placed within 15 cm in front of all speakers' mouths. The EGG system was not available for remaining subjects. Speakers first read the number of the order for each token and then paused before reading each target token on the reading list. Speakers took a break and drank water after every 100 tokens. The recording lasted for about 40 minutes.

## 2.5. Data analysis

The closed quotients of the glottal wave form were calculated with a hybrid method (CQ\_H) [8]. Following previous studies [10, 11],  $H1^*-H2^*$  amplitude was calculated with VoiceSauce [21]. Corrected  $H1^*-H2^*$  values were used, because the vowel quality was not controlled. A negative  $H1^*-H2^*$  value indicate creaky voice quality. The closed

quotient,  $f_0$ , and  $H1^*-H2^*$  values at every 10% in time during the vowel were extracted and analyzed with one-way ANOVA (checked vs. unchecked tones).



### 3. Results

#### 3.1. F0

As shown in Figure 1, the f0 contours of tones 5, 3, 53, and 3 fell through out the entire vowels. The f0 onset of unchecked tone 53 was the highest, followed by tone 3, then tone 5, and finally tone 31.

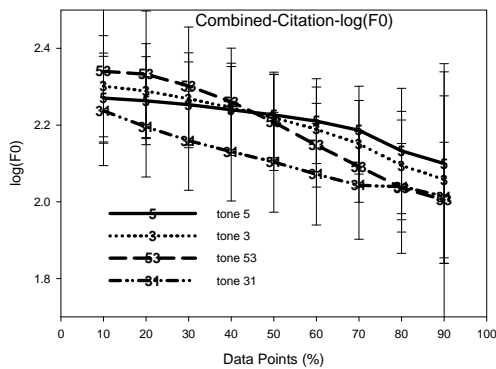


Figure 1: F0 contours.

#### 3.2. EGG

As shown in Figure 2, there was a significant effect of checked tones vs. unchecked tones on closed quotients produced by middle aged female speaker, CSG, from Northern Quanzhou region ( $F(1, 302) = -3.74, p < .01$ ), but not any other speakers. The closed quotient may not be an effective parameter in documenting voice quality. Other EGG parameters should be used.

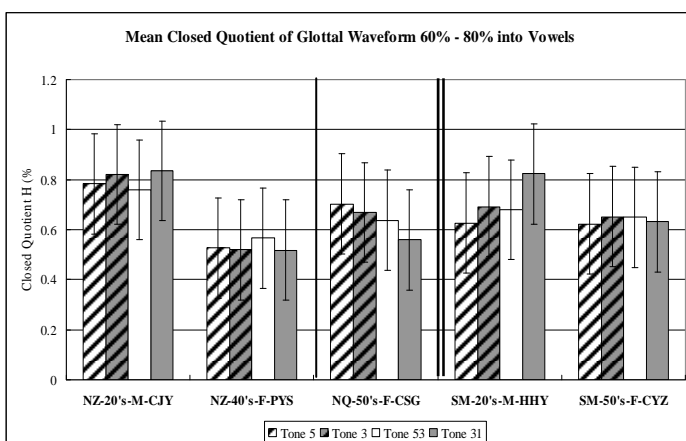


Figure 2: Close quotients of glottal wave form.

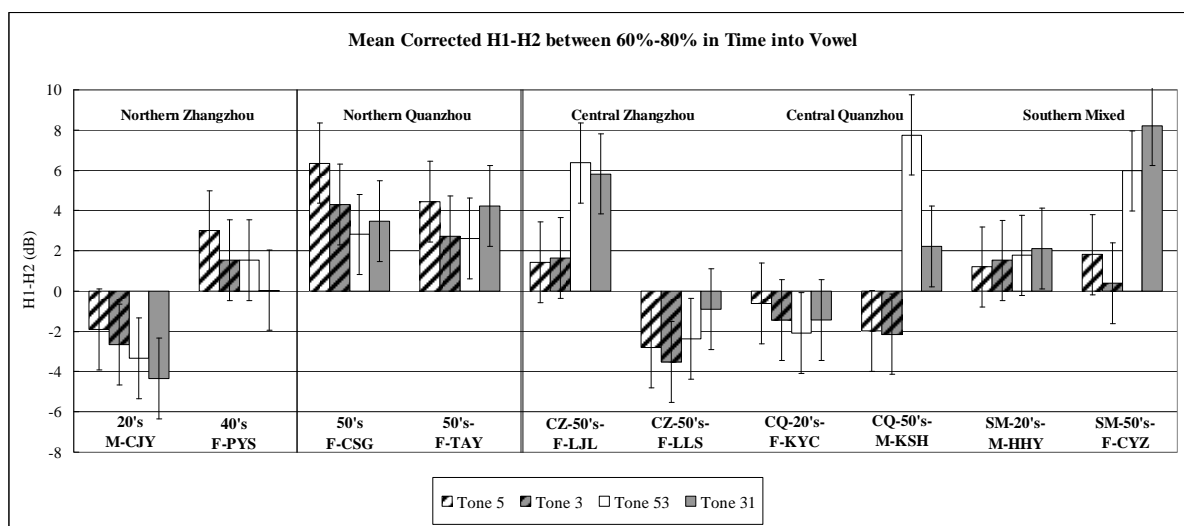
#### 3.3. Spectral tilt: H1\*-H2\* corrected

There were significant effects of tones on spectral tilt, H1\*-H2\*, produced by a young speaker, CJY, from Northern Zhangzhou region ( $F(1, 327) = -3.08, p < .01$ ), a middle aged female speaker, CSG, from Northern Quanzhou region ( $F(1, 310) = -4.6, p < .01$ ), two female middle aged speakers, LJL and LLS, from Central Zhangzhou region (LJL:  $F(1, 298) = 7.9, p < .01$ ; LLS:  $F(1, 298) = 4.98, p < .01$ ), a middle aged male speaker, KSH, from Central Quanzhou region ( $F(1, 312) = 16.99, p < .01$ ), and a middle aged female speaker, CYZ, from Southern Mixed region ( $F(1, 327) = 11.854, p < .01$ ).

In other words, excluding one young speaker, CJY, significant effects of tone on H1\*-H2\* were only observed on spectral tilt data from middle aged speakers.

As shown in Figure 3, in the two northern speakers' CSG and CJY, data, the H1\*-H2\* values of checked tones were higher than unchecked tones.

For these two northern speakers, the voice quality of checked tones was NOT more tense / creaky than unchecked tones as suggested by fiber optical or laryngoscopic studies. However, in central and southern middle aged speakers' productions, the H1\*-H2\* of checked tones were smaller in values than



**Figure 3:** Spectral tilt.

uncheck tones. Checked tones were tenser/ creakier than unchecked tones in central and southern speakers' data.

Even among and central and southern middle aged speakers, only in KSH's spectral tilt data was there a creaky versus modal versus breathy voice quality distinction. KSH produced checked tones with negative H1\*-H2\* values, indicating a creaky voice quality. He produced unchecked tones with positive H1\*-H2\* values, indicating a modal / breathy voice quality.

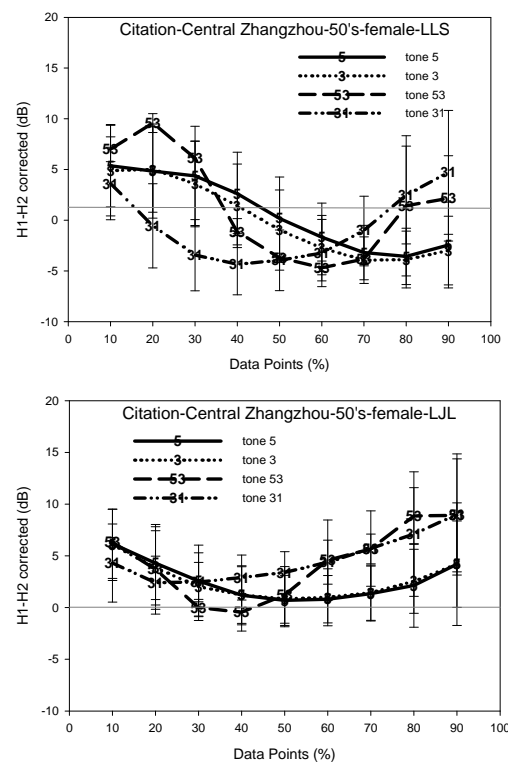
### 3.4. Timecourse

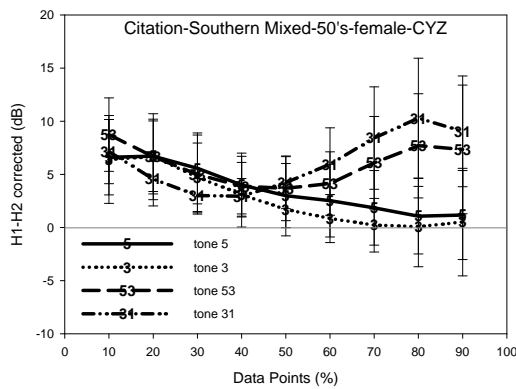
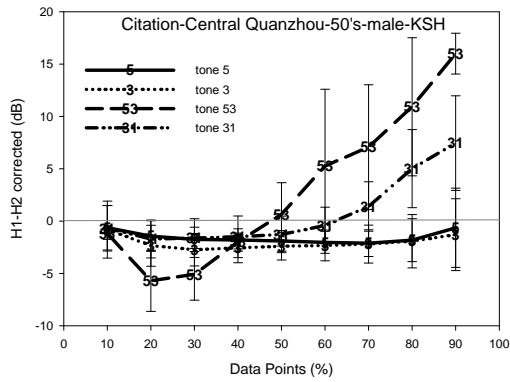
Among the central and southern speakers who produced checked tones with a tenser / creakier voice quality, the voice quality contrast between checked and unchecked tones was either in the middle or final portion of the vowels. As shown in Figure 4, the spectral tilt

distinction between checked tones, 5 and 3, and unchecked tones, 53 and 31, started after

40% in time into the vowels produced by LLS, after 50% in time for KSH and CYZ, and after 60% in time for LJJ. The acoustic cues for distinguishing checked tones from voice quality distinction occur in the final half of the vowels.

**Figure 4:** Times course of spectral tilts.





### 3.5. Prosodic effect

As shown in Table 2, tense phonation can be robustly observed among monosyllabic morphemes produced in citation form but the tense phonation contrast gradually disappeared in disyllabic word. CPP is the only parameter that maintain a clear contrast between tense phonation of checked tones and modal phonation of unchecked tones. As the phonation contrast started to disappear, the closure duration of final /p, t, k/ can be clearly observed before the onset of next syllable, as shown in Figure 5.

Table 2. Tense phonation in citation form and disyllabic words.

Citation	CQ_H	PIC	H1-H2	B1	CPP
>40	NZ	F, M	F, M	F	F, M
	SM	F, M	F, M	F, M	F, M
<30	NZ			F	F, M
	SM	M	M	F	F, M
<b>Word</b>					
>40	NZ		F		M
	SM		F		F, M
<30	NZ				M
	SM	F			M

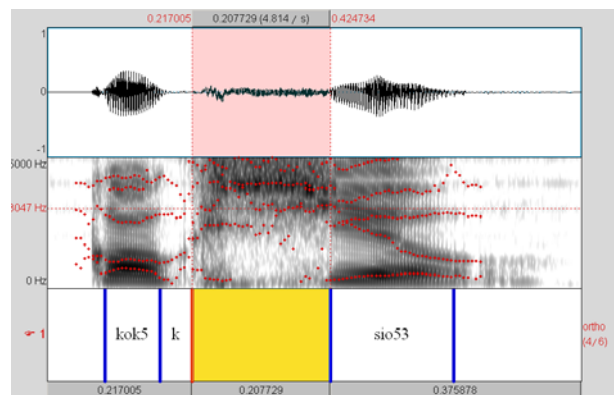


Figure 5: Duration of final stops in disyllabic words .

## 4. Discussion

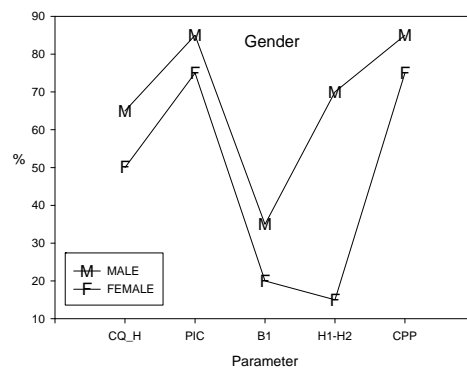
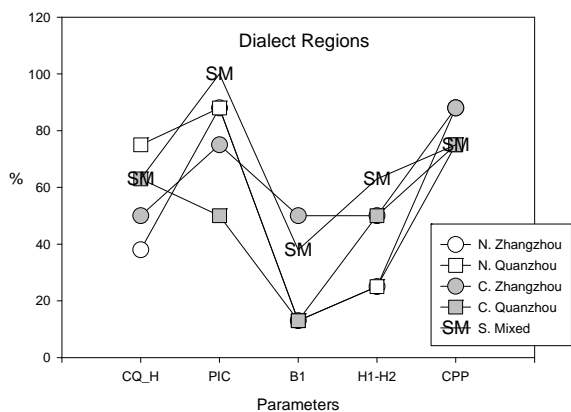


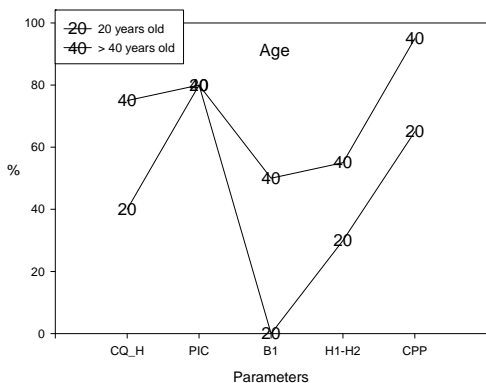
Figure 6: Effect of Gender on voice quality

As shown in Figure 6, among the 20 male and 20 female speakers, Male speakers showed a more tense voice quality than females. The contact quotient were longer, PIC higher, B1 narrower, H1-H2 lower and CPP higher for male speakers.



**Figure 7: Effect of dialect region on voice quality**

As shown in Figure 7, there is an effect of dialect region on voice quality. Speakers from southern mixed dialect region showed a tenser voice quality than speakers from other regions. The contact quotient were longer, PIC higher, B1 narrower, H1-H2 lower and CPP higher for speakers from southern mixed dialect region.



**Figure 8: Effect of Age on voice quality**

As shown in Figure 8, speakers above 40 years old showed a tenser voice quality than speakers under 30 years old. The contact quotient were longer, PIC higher, B1 narrower, H1-H2 lower and CPP higher for speakers above 40 years old.

The closed quotient data used here is not effective in distinguishing the voice quality of checked tones from unchecked tones. Further study on other EGG parameters, such as the beginning of contact phase, the end of contact phase, and the differences between minimum or maximum amplitude values in each cycle of EGG should be conducted.

Younger speakers do not maintain a voice quality distinction between checked and unchecked tones. This could be attributed to the influence of Mandarin, Taiwan's office language. Mandarin does not have checked syllables and so there is no need for such a distinction. Middle aged speakers whose native language is Taiwan Min still maintain the quality distinction. More data from different regions, ages and both genders should be collected to reveal the progress of checked tone voice quality changes in Taiwan.

## 5. Synegetic activities

Part of this project was conducted in the Phonetic Lab at University of California Los Angeles in collaboration of the NSF grant

BCS-0720304 “Linguistic uses of phonation across languages” project, for 2007 – 2011, with Prof. Pat Keating in Linguistic Department, Prof. Abeer Alwan in Electronic Engineering Department, Prof. Jody Kreiman in Head and Neck Surgery Department, Prof. Christina Esposito of Macalester College. , Jody Kreiman, and Abeer Alwan.

(<http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=0720304>)

Following is the abstract of the NSF project “Voice quality is due in part to patterns of vibration of a speaker's vocal folds inside the larynx. In some languages different voice qualities can distinguish word meanings, so each speaker must control multiple patterns of vibration. This project studies how speakers of different languages produce such voice contrasts when speaking, and how they perceive them when listening. Physiological recordings, acoustic measurements, and perceptual responses are analyzed to uncover the overall multi-dimensional phonetic space for voice, and to determine which subpart of that overall space each language uses.

The interdisciplinary research team for the project offers a unique approach to understanding linguistic voice quality, combining linguistic phonetics, voice science, and electrical engineering. Semi-automation of acoustic measurements of voice samples, with

new algorithms for recovering voice properties from standard audio recordings, allows large-scale data collection and analysis. These tools will be made available publicly for use by other researchers. The cross-language perspective of the project takes advantage of the many languages spoken in the Los Angeles area. The results of the project will be relevant for (1) describing languages, including documenting endangered languages; (2) high-quality synthesis of a wide range of normal voice qualities; (3) use of voice information in automatic speech recognition; (4) speaker (voice) recognition; (5) sociolinguistic study of voice quality variation as a function of linguistic/regional/social groups, including in bilingual populations; (6) study of voice quality variation with affect, emotion, etc.; (7) second-language learning of language-specific voice quality.”

## 6. Acknowledgement

This project was supported by National Science Council in Taiwan (NSC 98-2401-H-009-040-MY3).

## 7. REFERENCES

- [1] Abramson, A., L-Thongkum, T., Nye, P. W. 2004. Voice register in Suai (Kuai): An Analysis of perceptual and acoustic data, *Phonetica* 61, 147-171.
- [2] Andruski, J., Ratliff, M. 2000. Phonation types in production of phonological tone:

- the case of Green Mong, *J. of the International Phonetic Association*, 30(1/2), 37-61.
- [3] Andruski, J. E. 2006. Tone clarity in mixed pitch / phonation-type tones. *J. of Phonetics*, 34, 388-404.
- [4] Ang, U.-j. 2003. *The Motivation and Direction of Sound Change: On the Competition of Minnan Dialects Chang-chou and Chuan-chou and the Emergence of General Taiwanese*. Unpublished PhD. Dissertation, National Tsing Hua University TAIWAN.
- [5] Ang, U.-j. 2009. On the dialect typologies and isoglosses of Taiwan Min. *J. of Taiwanese Languages and Literature*, 3, 239-309.
- [6] Chen, S.-c. 2009. The vowel system change and the Yin-/Yang-entering tonal variations in Taiwanese Hokkien. *J. of Taiwanese Languages and Literature*. 3, 157-178.
- [7] Cheng, R. 1973. Some notes on toe sandhi in Taiwanese. *Linguistics*, 100, 5-25.
- [8] EggWorks. <http://www.linguistics.ucla.edu/faciliti/facilities/physiology/egg.htm>, 2010.
- [9] Edmondson, J. A., Chang, Y.-c., Huang, H.-c., Hsieh, F.-f., Peng, Y. 2010. Reinforcing voiceless stop coda in Taiwanese, Vietnamese and other east and southeast Asian languages: Laryngoscopic case studies. LabPhon 12. New Mexico
- [10] Esposito, C. 2006. *The Effects of Linguistic Experience on the Perception of Phonation*. UCLA Ph.D. dissertation.
- [11] Garellek, M., Keating, P. 2010. The acoustic consequences of phonation and tone interactions in Jalapa Mazatec. *UCLA Working Papers in Phonetics*, 108, 141-163.
- [12] Huffman, M. K. 1985. Measures of phonation type in Hmong. *UCLA Working Papers in Phonetics*, 61, 1-25.
- [13] Huffman, M. K. 1987. Measures of phonation type in Hmong. *J. Acoust. Soc. Am.* 81(2), 495-504.
- [14] Iwata, R., Sawashima, M., Hirose, H. 1981. Laryngeal adjustments for syllable-final stops in Cantonese. *Annual Bulletin Research Institute Logopediatrics and Phoniatics*. 15, 45-54.
- [15] Javkin, H. R., Maddieson, I. 1983. An inverse filtering study of Burmese creaky voice. *UCLA Working Papers in Phonetics*, 57, 115-125.
- [16] Keating, P., Esposito, C., Garellek, M., Khan, S., Kuang, J. 2010. Phonation contrast across languages. *UCLA Working Papers in Phonetics*, 108, 188-202.
- [17] Maddieson, I., Ladefoged, P. 1985. "Tense" and "lax" in four minority

languages of China. *UCLA Working Papers in Phonetics*, 60, 59-83.

- [18] Michaud, A. 2004. Final consonants and glottalization: New perspective from Hanoi Vietnamese, *Phonetica*, 61, 119-146.
- [19] Pan, H.-h. 1991. Voice quality of Amoy falling tones. *J. Acoust Soc Am.* 90, 2345.
- [20] Pan, H.-h. 2005. Voice Quality of Falling Tones in Taiwan Min. *Proc. Interspeech 2005 Lisboa*, 1401-1404.
- [21] Shue, Y.-L., Keating P., Vicenik, C. 2009. VoiceSauce: A program for voice analysis. *J. Acoust Soc Am.*, 126, 2221.