



自然語料中語音變異對語音認知之影響—— 語音語料庫之建置及其應用

馮怡藜*

一、計畫緣起

人類使用之語類可略分為三大項，即閱讀語類、預演語類與自然語類。閱讀語類通常出現於朗讀或念稿，語法與文本最相當，語音也與字典音最近似；預演語類則如廣播、電視訪問等，雖非照稿念，說者卻因說話內容經事前準備，語法與文本的相似度仍高，而語音雖不完全如字典音，卻也相去不遠；自然語類則是我們日常生活裡和朋友聊天、去商家購物等所使用之自發性語言，由於未經事先撰稿排練，語法與文本相距最遠，語音也與字典音相差最多，變化最大。大體而言，一般人除非離群索居，一天當中最常使用的語類，當屬自然語類。

然而，弔詭的是，語音學最常用來做研究的語類，卻是最不常使用的閱讀語類。主要原因在於，語音學基於實驗科學的要求，需要控制不必要的變因，且由於聲學分析的關係，錄音品質需相當乾淨，因此對帶有太多不可控制因素，又很難於錄音間環境中取得之自發性語料一向很少涉獵。當然，這樣的結果是相當可惜的，畢竟以大量閱讀語料所取得的研究成果，竟無法因建立語言常模而一窺口語現象的全貌，實屬遺憾。

爲了彌補這項缺失，近幾年來，有愈來愈多的語音學研究，嘗試利用自然語料進行。由於自然語料中單一現象的密度偏低，不利研究進行，故目前的趨勢乃以建立自然語音語料庫爲主，利用錄音間的專業錄音設備，取得大量乾淨的錄音資料後，將語料予以適當後製標記處理，以利之後各類研究議題之進行。目前英語 Buckeye 自然語料庫 (Raymond, et al., 2001)、日語 CSJ 自然語料庫 (Maekawa, Koiso, Furui, & Isahara, 2000)、德語 Kiel 自然語料庫

* 國立臺灣大學語言學暨腦與心智科學研究所副教授

(Kohler, 1996) 及中研院現代漢語對話語音語料庫 (Tseng, 2005) 皆以此種方式進行。

由於本人一向對自然語體中多變的語音現象感到相當著迷，因此本計畫針對臺灣特有的華閩雙語人口進行有系統的語料蒐集，並利用所建置的語料庫，探討臺灣華語舌尖前後音之變異現象，及此項變異對語音感知所造成的影響。本計畫共分三項實驗，分三年進行，以下簡述三年來的研究成果。

二、實驗一：臺灣華語自然語音語料庫之建置

實驗一主要是建置北高兩地之自然語音語料庫，總共錄製 15 位發音人，其中臺北地區四男三女，高雄地區四男四女。所有發音人皆為華閩雙語人士，錄音時年齡皆為 20 至 35 歲，且 3 至 18 歲未曾離開出生地超過六個月以上。錄音方式比照英語 Buckeye 自然語料庫的方式，採用訪談形式，訪談內容以中性題材為主，時間約半小時左右。表一為語料庫相關資訊。語料庫共有近八小時、超過十萬字的錄音。發音人以利克特量表自評華閩語流利度，以 1 為最不流利，7 為最流利。

所有錄音檔皆經漢字與漢語拼音轉寫，再將漢語拼音的結果，以 Praat (Boersma & Weenink, 2009) 上標記。漢字轉寫內容經由中文斷詞系統斷詞，並加上詞類標記 (Chen, 1998)。由於斷詞系統原為文本而設，而本語料庫含許多口語用詞，故時有斷詞系統無法處理的情況。因此經斷詞系統作初步處理後，所有詞類皆經人工修正確認。再者，由於中文詞的界定相當不易，斷詞原則上採 Huang, Chen, Chen, Wei & Chan (1996) 所訂的標準，並增加「語篇標記詞」一類，以符合語料庫含大量此詞類的需求。表二為語料庫所標記的詞類。

表一：語料庫相關資訊。M：平均；T：總數

區域	發音人	年齡	錄音時長	音節數	華語流利度	閩語流利度
臺北男性	CZX	24	30m10s	8641	7	6
	HSK	33	33m09s	8285	7	6
	JXW	35	30m27s	5379	5	7
	YYS	28	31m53s	7676	6	6
臺北女性	LRL	22	30m20s	6671	7	4
	XHR	31	30m53s	8111	7	7
	XHY	32	32m13s	8243	7	6

(續下頁)



表一：語料庫相關資訊。M：平均；T：總數（接上頁）

區域	發音人	年齡	錄音時長	音節數	華語流利度	閩語流利度
高雄男性	GWH	21	30m14s	6486	7	4
	KCZX	23	30m37s	7082	6	7
	LJH	24	30m37s	7550	7	5
	LZW	27	30m32s	7044	6	7
高雄女性	CJH	22	30m43s	4823	7	5
	CN	20	30m37s	7151	7	4
	GYX	22	35m08s	8635	6	5
	WJL	35	31m19s	8054	7	7
統計		M = 26.6	T = 468m52s	T = 109831	M = 6.6	M = 5.73

表二：語料庫標記詞類

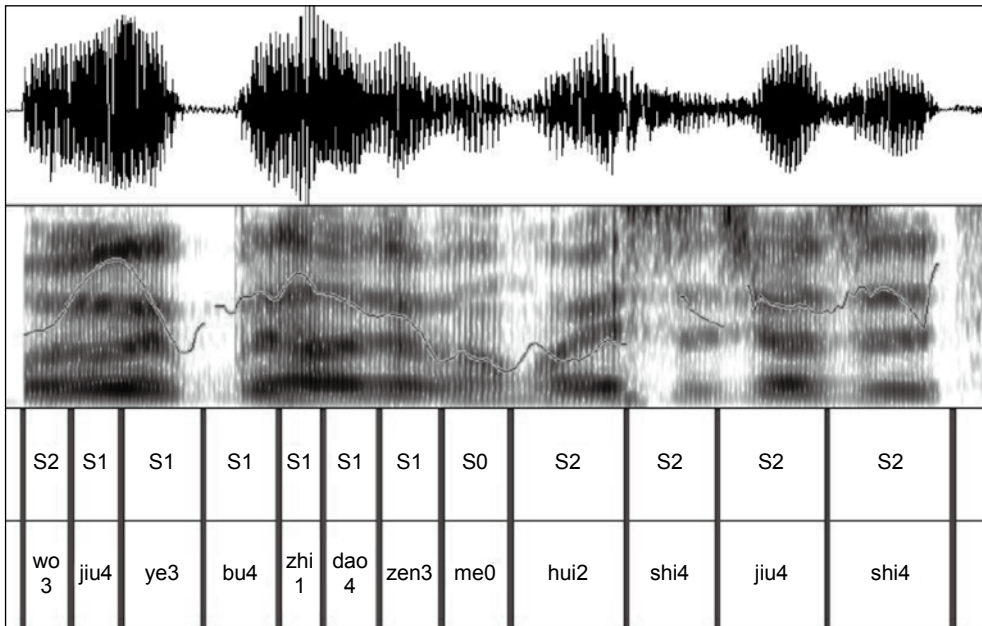
實詞	舉例	虛詞	舉例
名詞	「桌子」	代名詞	「我」
動詞	「吃飯」	指代定詞	「這」
形容詞	「主要」	DE	「的」
副詞	「其實」	SHI	「是」
量詞	「種」	介詞	「在」
數量定詞	「很少」	後置詞	「之類」
數詞	「三」	連接詞	「所以」
		否定詞	「不」
		時態標記	「著」
		語篇標記詞	「就是說」

除詞性標記外，語料庫尚有重音標記。重音標記主要依 Pan-Mandarin ToBI (Peng, et al., 2005) 韻律標記系統的原則進行。ToBI 標記乃依閱讀語料建立，而閱讀語料中重音層次與聲調調型的完整性呈高度相關，因此原本 ToBI 的重音定義主要著重聲調調型的呈現，調型愈完整，代表重音層次愈高。然而，這樣的定義在自然語料卻無法完全適用，主要原因在於自然語料中的聲調調型往往並不完整，故若依原本 ToBI 的重音定義，幾乎完全無法有高層次重音的標記，與聽者感知並不相符。因此在進行重音標記時，我們保留 ToBI 重音的四個層次，但根據自然語料的特性將原本一些定義做適當的修正，以符合語料的特性，修正後定義如表三所示。除原有的調型定義外，增加音強、音長與音段特性的定義，以期能讓自然語料中的重音層次標記更符合實際需求。圖一為 S0、S1 及 S2 之標記例句。由該圖可知，標記 S1 之音

節不僅調型多半無法維持，且音段訊號也有相當程度的央化，發音目標並未達成。圖二則為 S2 及 S3 之標記例句。由該圖可知標記 S3 之音節「辣」不僅音強較周遭音節高出許多，與句中其他四聲字相較（如「印度」），該字的調閥擴大，且音長較預設值延長許多，音段訊號也相當完整，可知說者相當強調此一音節。

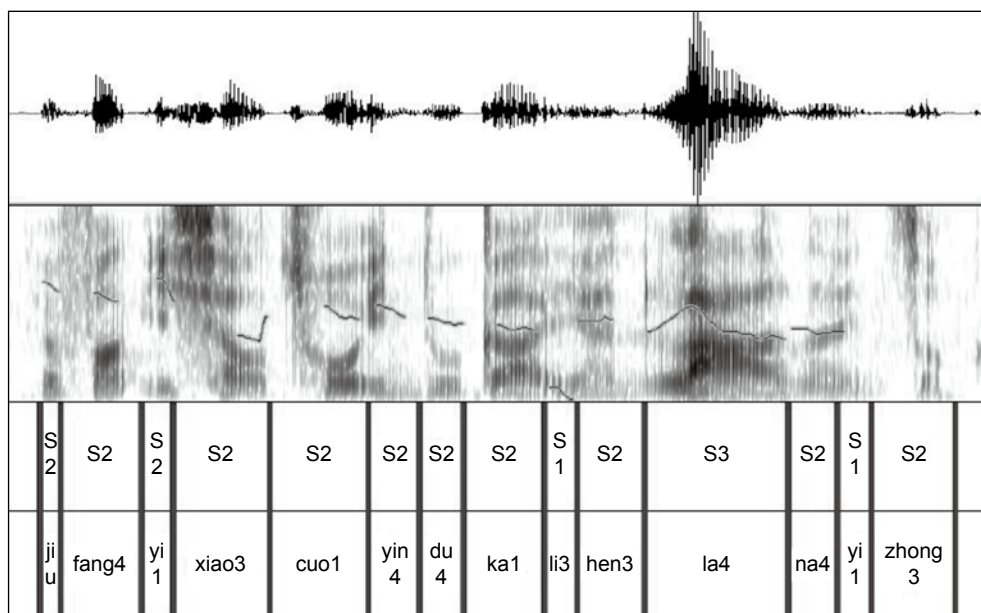
表三：重音層次修正後定義

重音層次	調型	音強	音長	音段特性
S3	調閥擴大 / 上提	強	延長	發音目標完全達成
S2	預設調型	預設值	預設值	發音目標大部分達成
S1	調型輕化	弱	縮短	發音目標未達成
S0	輕聲調型	弱	縮短	發音目標未達成



圖一：以「我就也不知道怎麼會是就是...」一句示範 S0、S1 及 S2 三個重音層次。標記第一層為重音標記，第二層為漢語拼音音節標記。

為確保修正後之重音定義之信度，我們先將語料庫中 10% 的資料交由兩位具 ToBI 訓練背景之語音學專家，在給予她們修正後之重音定義後，請她們進行獨立標記作業。結果顯示，二人的標記有 82% 相同，顯示此一標記定義具一定信度。因此，語料庫中所有的重音標記皆採用修正後之定義。



圖二：以「就放一小撮印度咖哩很辣那一種...」一句示範 S2 及 S3 二個重音層次。標記第一層為重音標記，第二層為漢語拼音音節標記。

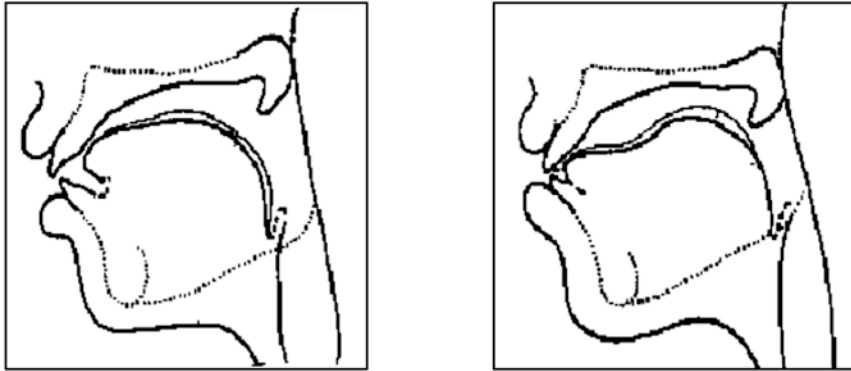
三、實驗二：臺灣華語舌尖前後音之變異現象

實驗二主要探討臺灣華語舌尖前後音於自然語體中的變異現象。華語共有三對舌尖前後清音，如表四所示。根據 Ladefoged & Wu (1984) 的研究，舌尖前後音的發音差異並不在捲舌與否，二者皆由舌面接近齒齦發音而成，此與一般認知相左。X 光研究結果顯示二者的差異點主要有三（圖三）：其一、舌尖前音發音部位較前，於齒齦前，而舌尖後音發音較後，於齒齦上；其二、舌尖後音有舌下腔，而舌尖前音則無；其三、舌尖前音的舌面有輕微凹陷，而舌尖後音則無。

表四：華語三對舌尖前後清音

	擦音	不送氣塞擦音	送氣塞擦音
舌尖前音	[s]	[ts]	[tsʰ]
舌尖後音	[ʃ]	[tʃ]	[tʃʰ]

之前聲學研究顯示，擦音分析主要以頻譜動差為主，包含頻譜重心頻率 (M1)、標準差 (M2)、偏態 (M3) 及峰度 (M4)。M1 訊號在分辨噝音上尤其



圖三：Ladefoged & Wu (1984) 的 X 光發音研究。左圖為華語的 [ʃ]，右圖為 [ʃ]。

強健 (Jongman, Wayland, & Wong, 2000)，與前口腔長度成反比，亦即前口腔愈短，高頻能量愈強，而 M1 值則愈高。Li, F. (2008) 發現以擦音部分中間 40 微秒所計算出的 M1 值可將中國華語的三組舌尖前後清音清楚區分，Jeng (2006) 於臺灣華語亦發現類似情形，且擦音時長於舌尖後音較短。

雖然一般對臺灣華語的印象是舌尖前後音合流，Tse (1998) 量測二者的功率譜極小值後發現，年輕華語使用者多半仍能區分二者。Jeng (2006) 則發現，合流現象依語境而異，於問答時較多，於朗讀時較少。Lin (1983) 另顯示臺灣年輕女性區分舌尖前後音的能力較男性強，但此種現象僅及於正式語體 (Li, M.-C., 1995)。於非正式語體，女性合流現象較男性常見。Lin (1983) 亦發現，青少年期前住在臺北的受試者較不住在臺北者更易保有舌尖前後音的分別，顯示地域會有一定程度的影響。雖然整體而言，合流現象主要以舌尖後音改為舌尖前音呈現，但仍有相反的狀況，即舌尖前音改以舌尖後音呈現 (Li, M.-C., 1995; Lin, 1983)，Chung (2006) 將此種現象歸因於矯枉過正。

由於之前的研究少有以自然語料為主要研究素材，本實驗即針對臺灣華語自然語料中舌尖前後音的呈現變異，進行一系列分析。由於許多語言中重音 (Beckman & Edwards, 1994) 與詞性 (Raymond, Dautricourt, & Hume, 2006; van Bergem, 1993) 皆與音段的發音有一定的相關性，故變因除已知的性別與地域等外部因素外，亦包括重音與詞性等內部因素，以期能對臺灣華語之舌尖前後音的變異，有更全面性的了解。

本實驗首先於實驗一所錄製的語料庫中找出所有舌尖前後清音，並標記其擦音部分。聲學量測的方式主要參考 Li, F. (2008)，不過由於自然語料語



速較快，擦音部分較短，頻譜切片改為 10 微秒，且因舌尖前後音於 2000 赫茲以下的頻譜差別不大，高通濾波器的截止頻設在 2000 赫茲，以濾掉低頻雜訊。此外，之前研究顯示 M1 訊號於分辨舌尖前後音相當強健 (Jeng, 2006; Li, F., 2008)，故將濾過的頻譜取其 M1，以查看舌尖前後音是否合流。

整個語料庫共含 20824 個舌尖前後清音，其中 16258 個有效。有效個數中，有 74 個舌尖前音聽覺上確定發為舌尖後音，有 85 個舌尖後音聽覺上確定發為舌尖前音。其餘 16099 個樣本包含照字典發音及聽覺上無法確認是否為合流之音（之後本文以「字典音」概括稱之），為本計畫主要分析範圍，詳細分布如表五。當中，符合字典音的字於詞性與重音層次之分布並不平均（表六），故無法同時分析所有重音層次，因此以下就足夠數量之組合，進行系統性分析。

表五：舌尖前後清音之發音分布

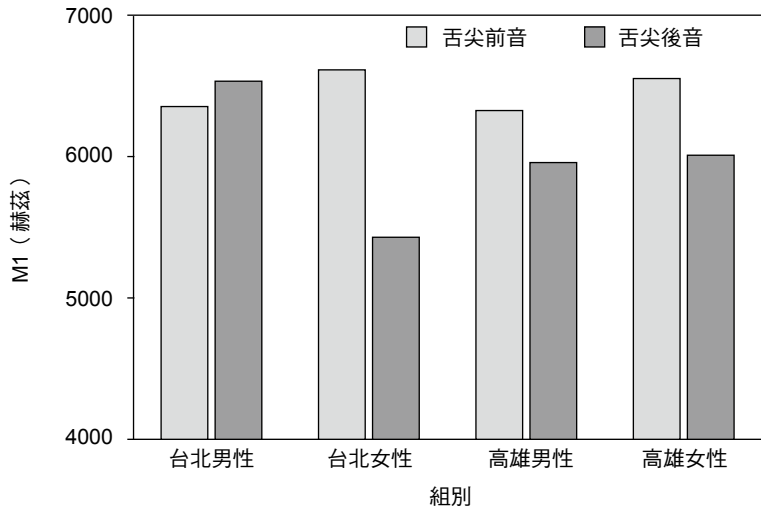
		實際發音		總計
		舌尖前音	舌尖後音	
字典音	舌尖前音	4007	74	4081
	舌尖後音	85	12092	12177
總計		4092	12116	16258

表六：照字典音之舌尖前後清音與重音層次及詞性之分布

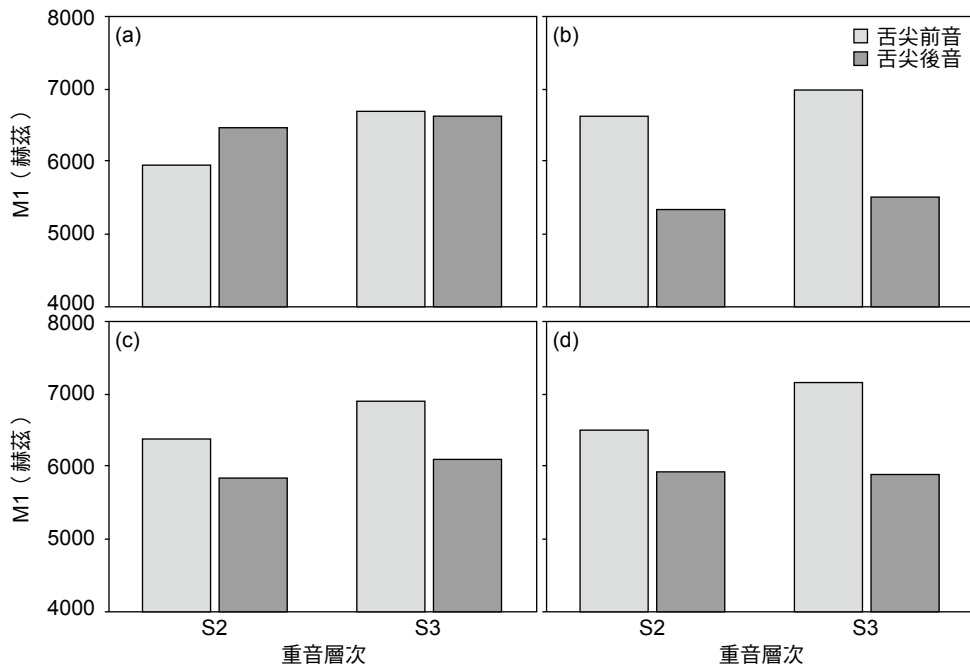
	實詞		虛詞		總計
	舌尖前音	舌尖後音	舌尖前音	舌尖後音	
S0	109	0	66	16	191
S1	11	68	7	110	196
S2	2274	6444	962	3970	13650
S3	544	1398	34	86	2062
總計	2938	7910	1069	4182	16099

圖四為臺北與高雄受試者之平均 M1 值。由於 S0 與 S1 的個數分布相對較不平均也較少，因此此圖僅包含 S2 與 S3 的資料。整體而言，臺北女性的舌尖前後音差異最大，臺北男性則幾無差異，高雄受試者介於二者之間，且男女的差異不大。

圖五為重音層次對舌尖前後音之發音影響。由表六可知，重音層次於實詞與虛詞之分布極不平均，因此僅就實詞於 S2 與 S3 之發音進行分析。結果



圖四：臺北與高雄各組受試者於 S2 與 S3 重音層次之舌尖前後音平均 M1 值。



圖五：重音層次對舌尖前後音平均 M1 值於 (a) 臺北男性、(b) 臺北女性、(c) 高雄男性與 (d) 高雄女性之影響。

顯示，除高雄女性外，其他組受試者由 S2 到 S3 皆呈現前音位化的現象，且以舌尖前音的程度較高。高雄女性表現重音的策略則略有不同，只有舌尖前

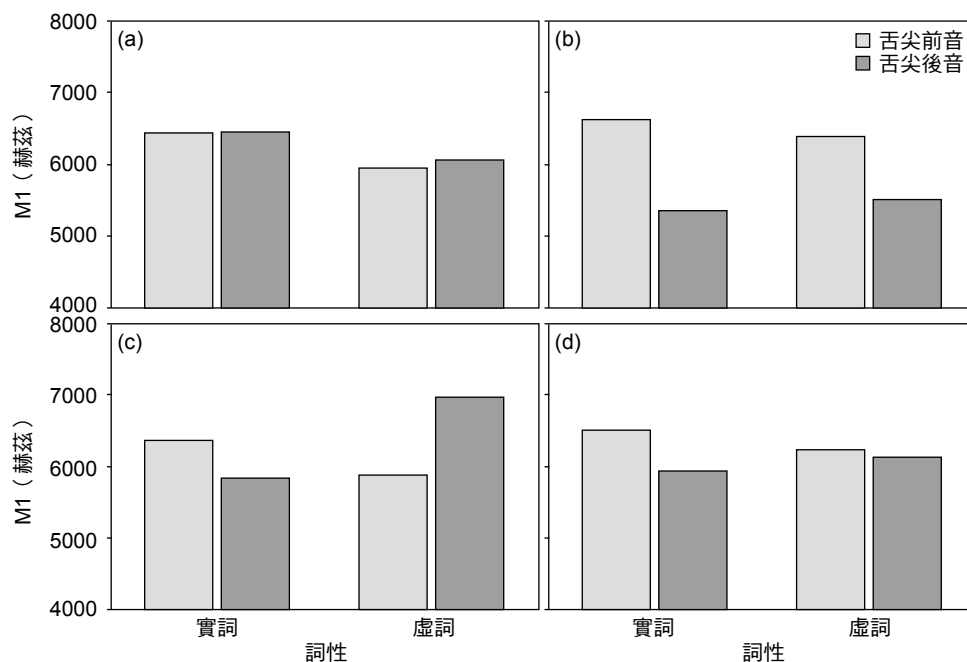


音呈現前音位化的現象，舌尖後音則無。受前音位化影響，除臺北男性因舌位無明顯差異外，其餘各組受試者由 S2 到 S3 皆呈現舌尖前後音差異變大的現象。

圖六為詞性對舌尖前後音之發音影響。由表六可知，虛詞於舌尖前後音之分布不平均，故僅就重音層次為 S2 之資料進行分析。結果顯示，除臺北男性外，其餘各組受試者於虛詞時舌尖前音皆有後音位化之趨勢，而舌尖後音則有前音位化之趨勢。臺北男性不論舌尖前後音，皆有後音位化之趨勢。而受詞性影響，女性受試者於虛詞時舌尖前後音差異變小。

表七為各組受試者舌尖前後音合流之分布情形。整體而言，高雄女性較易出現舌尖前音發為舌尖後音的情形，而舌尖後音發為舌尖前音的情形，除高雄男性較少外，其餘各組受試者出現個數相差不大。

實驗二的結果顯示，臺灣華語舌尖前後音的變異現象相當複雜，有少部分呈現完全合流的情況，可由聽覺上直接做判斷，但大部分則相當程度受到韻律相關因素而呈現不同合流程度。外部因素如地區與性別亦有不小的影響



圖六：詞性對舌尖前後音平均 M1 值於 (a) 臺北男性、(b) 臺北女性、(c) 高雄男性與 (d) 高雄女性之影響。

表七：各組受試者舌尖前後音合流之分布

	臺北男性	臺北女性	高雄男性	高雄女性	總計
舌尖前音→舌尖後音	12	5	4	53	74
舌尖後音→舌尖前音	26	25	9	25	85
總計	38	30	13	78	159

力，整體而言，臺北男性合流的情況最為明顯，除少數語境外，其舌尖前後音幾無明顯區分，而臺北女性舌尖前後音之差異最大，幾無合流現象。高雄受試者則介於二者之間。

四、實驗三：臺灣華語舌尖前後音變異之感知現象

由之前研究與實驗二的結果顯示，臺灣華語舌尖前後音的變異現象隨著不同族群、性別、韻律語境等而有不同程度的合流現象，因此實驗三的目的主要在了解此種變異現象對聽者感知的影響，以了解變異音是否已成為聽者感知之音韻規則。

本實驗採用跨通道形式促發詞彙判斷作業進行，實驗材料包括促發物與目標物。受試者先聽到促發物，接著在螢幕上看到目標物，且需針對目標物在反應箱的按鍵上做詞彙判斷。實驗刺激之目標物分三大組：第一組為 15 個以舌尖前音為首之雙音節詞（如「思考」），第二組為 15 個以舌尖後音為首之雙音節詞（如「失敗」），第三組則為 15 個以不送氣塞音為首之雙音節詞（如「個性」）。前兩組為實驗組，後一組為控制組，三組刺激詞頻相當。每個目標物皆有三種促發物：正確音、變異音及非相關音。正確音即字典發音；變異音三組目標物設定有別：舌尖前後音互為變異音，而不送氣塞音之變異音則設定為送氣音；非相關音為控制組。表八為實驗材料之一例。

受試者共 45 人，分三組，皆為閩南籍臺灣華語母語人士，年齡為 18 至 25 歲。實驗中每組受試者皆對 45 個實驗刺激目標物做反應，但對每個目標物，每組人只會聽到一種促發物。所有受試者皆無聽力或語言障礙病史。

實驗結果分析分為正確率及反應時間。正確率如圖七所示，整體而言相當地高（>96%），但可看出不同促發物的影響。正確音與變異音之正確率明顯高於非相關音，顯示前二者皆可順利促發目標物，而後者促發效果較差。反應時間之結果如圖八所示，不同促發物有顯著影響，不論是何種目標物，

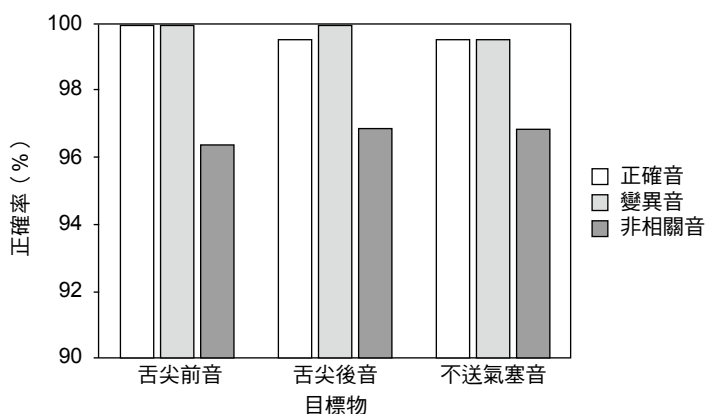


非相關音的促發速度明顯較慢。對含舌尖前後音之目標物而言，正確音與變異音之促發效果相當，但對含不送氣塞音之目標物而言，變異音之促發速度較正確音慢。

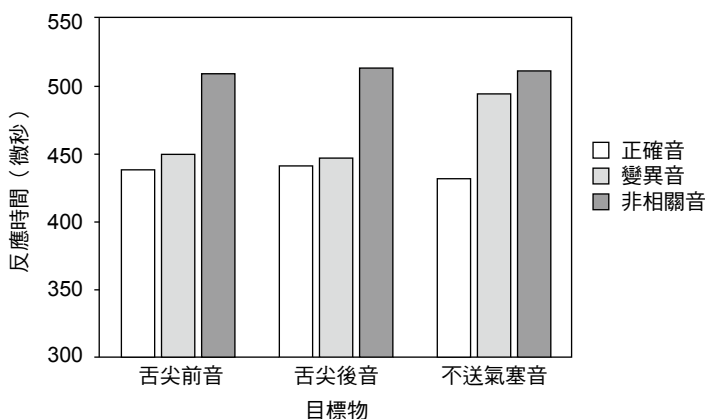
表八：實驗三材料促發物及目標物之例子

促發物(聽)	舌尖前音	舌尖後音	不送氣塞音
正確音	[sɿ] 考	[tʂʰɑʊ] 架	[kɿ] 性
變異音	[ʂɿ] 考	[tʂʰɑʊ] 架	[kʰɿ] 性
非相關音	房屋	訓練	游泳
目標物(看)	思考	吵架	個性

實驗三的結果顯示，臺灣華語使用者似乎將舌尖前後音互視為同位音，因此「[sɿ] 考」與「[ʂɿ] 考」皆可促發「思考」一詞，而「[tʂʰɑʊ] 架」與「[tʂʰɑʊ]



圖七：不同促發物下三組目標物之反應正確率。



圖八：不同促發物情境下對目標物之反應時間。

架」亦皆可促發「吵架」一詞，這意味舌尖前後音之變異現象已成為聽者感知之音韻規則，且正確音與變異音的促發力似乎有齊鼓相當之勢。

五、總結

在國科會人文處的補助下，本計畫得以順利進行，不僅完成臺灣北高兩地自然語音語料庫之建置，可提供日後更多相關研究的語料基礎，更藉由該語料庫探討臺灣華語中常見卻很難於自然語體中進行定量分析之舌尖前後音合流情形，使得我們對該現象的多層次表現，有更進一步的認識。由於考量到語音變異不僅是產製現象，更會對感知系統造成影響，因此本計畫以心理語言學實驗法探討臺灣華語使用者對此語音變異的感知表現，讓我們對於臺灣華語舌尖前後音的變異現象，由產製至感知有一更全面性的了解。

相關文獻

- Beckman, M. E. & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (Ed.), *Papers in Laboratory Phonology III: Phonological structure and phonetic form* (pp. 7-33).
- Boersma, P. & Weenink, D. (2009). Praat: Doing phonetics by computer (Version 5.1) [Computer Program].
- Chen, K.-J. (1998). Online Word Segmentation Aystem. Retrieved April, 2009, from <http://ckipsvr.iis.sinica.edu.tw/>
- Chung, K. S. (2006). Hypercorrection in Taiwan Mandarin. *Journal of Asian Pacific Communication*, 16 (2), 197-214.
- Huang, C.-R., Chen, K.-j., Chen, F.-Y., Wei, W.-C., & Chan, L.-I. (1996). A Segmentation Standard for Chinese Information Processing: Design criteria and content: Academia Sinica.
- Jeng, J.-Y. (2006). The acoustic spectral characteristics of retroflexed fricatives and affricates in Taiwan Mandarin. *Journal of Humanistic Studies*, 40 (1), 27-48.
- Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of Acoustical Society of America*, 108 (3), 1252-1263.
- Kohler, K., J. (1996). Labelled data bank of spoken standard German: the Kiel corpus of read/spontaneous speech. *Proceedings of the Fourth International Conference on Spoken Language Processing*, 1938-1941.
- Ladefoged, P. & Wu, Z. (1984). Places of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics*, 12, 267-278.
- Li, F. (2008). The phonetic development of voiceless sibilant fricatives in English, Japanese and Mandarin Chinese. Unpublished dissertation, The Ohio State University.
- Li, M.-C. (1995). A sociolinguistic variation study of Chinese retroflex initials / tʂ /, / tʂʰ / and / ʂ /



- in Taiwan. Unpublished thesis, Providence University, Taichung.
- Lin, C. C. (1983). A sociolinguistics study of the use of the retroflex sounds in Mandarin in college students in Taiwan. *Bulletin, College of Liberal Arts, National Central University*, 1, 1-15.
- Maekawa, K., Koiso, H., Furui, S., & Isahara, H. (2000). Spontaneous speech corpus of Japanese. *Proceedings of the 2nd International Conference on Language Resources and Evaluation*, 947-952.
- Peng, S.-h., Chan, M. K. M., Tseng, C.-y., Huang, T., Lee, O. J., & Beckman, E. M. (2005). Towards a pan-Mandarin system for prosodic transcription. In S.-A. Jun (Ed.), *Prosodic Typology: The phonology of intonation and phrasing* (pp. 230-270). New York: Oxford University Press.
- Raymond, W. D., Dautricourt, R., & Hume, E. (2006). Word-initial /t, d/ deletion in spontaneous speech: Modeling the effects of extra-linguistic, lexical, and phonological factors. *Language Variation and Change*, 18, 55-97.
- Raymond, W. D., Makashay, M. J., Dautricourt, R., Johnson, K., Hume, E., & Pitt, M. (2001). Variation in conversation: An introduction to the Buckeye Speech Corpus. Paper presented at the 140th Meeting of the Acoustical Society of America.
- Tse, J. K.-P. (1998). Taiwan diqu nianqingren zhi chi shi yu zi ci si zhen de bu fen ma [Do the young people of Taiwan really not distinguish between zh-, ch-, sh- and z-, c-, s-?]. *The World of Chinese Language*, 90, 1-7.
- Tseng, S.-C. (2005). Mandarin topic-oriented conversations. *International Journal of Computational Linguistics and Chinese Language Processing*. Special Issue: Annotated Speech Corpora, 10 (2), 201-218.
- van Bergem, D. R. (1993). Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication*, 12, 1-23.