Title

Lexical Effects in Phonemic Neutralization in Taiwan Mandarin

Journal

Proceedings of the Annual Meeting of the Berkeley Linguistics Society, 38(38)

ISSN

2377-1666

Author

Li, Ying-Shing

Publication Date

2014

Peer reviewed

Lexical Effects in Phonemic Neutralization in Taiwan Mandarin
Author(s): Ying-Shing Li
Proceedings of the 38th Annual Meeting of the Berkeley Linguistics Society (2014), pp. 307-327
General Session and Thematic Session on Language Contact
Editors: Kayla Carpenter, Oana David, Florian Lionnet, Christine Sheil, Tammy Stark, Vivian Wauters

Lexical Effects in Phonemic Neutralization in Taiwan Mandarin

Ying-Shing Li
Institute of Linguistics, Academia Sinica

1. Neutralization of Sibilant Onsets and Nasal Codas in Taiwan Mandarin

Colloquial Taiwan Mandarin has deviated from Guoyu [National language] or Standard Chinese in pronunciation, vocabulary, and even syntax. Such changes come from the linguistic contact with Taiwan Southern Min or natural diachronic linguistic drift (Kubler, 1985; Tung, 1994; Tsao, 2000). This new form of Taiwan Mandarin has become a lingua franca among speakers of the different backgrounds in Taiwan and a creole for new generations to acquire as their mother tongue (Her, 2009). One of the most noticeable segmental changes in Taiwan Mandarin is the merging of alveolar sibilants [ts, ts^h, s] and retroflex sibilants [tʂ, tʂ^h, ʂ]. The other one in Taiwan Mandarin is the neutralization of alveolar nasal coda [n] and velar nasal coda [ŋ].

Taiwan Mandarin consonant inventory consists of three sets of voiceless coronal sibilants, including unaspirated affricates, aspirated affricates, and fricatives (Chen, 1973). All of the sibilants occur only in the onset position. Among them, the alveolopalatal sibilants [tɕ, tɕ^{h,} ɕ] only precede the high medial vowels [i, y], whereas alveolar sibilants [ts, ts^h, s] and retroflex sibilants [tʂ, tʂ^h, ʂ] precede the other medial or nucleus vowels in common. Many previous studies on Taiwan Mandarin have found that the retroflex sibilants are approximating to the alveolar sibilants (Kubler, 1985; Wu, 1985; Li, 1986; Yao, 1987; Chen, 1991; Luo, 1991; Yeh, 1991; Rau and Li, 1994; Tung, 1994; Tse, 1998; Tsao, 2000; Chung, 2006). Tung (1994) explicitly illustrated six retroflex sibilant productions in Taiwan Mandarin whose fricative constrictions occurred variably in the coronal places of articulation, depending on the individuals' dialect backgrounds and the following vowels. Chung (2006) also described the variability of retroflex realizations in Taiwan Mandarin as "The degree of tongue (tip) retraction may vary considerably (as it does in Beijing Mandarin too, although a different range is covered), from highly retracted, through the palate-alveolar area [tʃ], [tʃ^h], [ʃ], all the way to dentals that are indistinguishable from the dental/apical *z*- [s̪], *c*- [ts̪^{h}], *s*- [s̪] series" (Chung, 2006:200).

Various studies suggested that the neutralization of retroflex sibilants and alveolar sibilants in Taiwan Mandarin was conditioned by multiple factors. Phonologically speaking, unaspirated affricated sibilants enhanced neutralization relative to the two other sibilants (Tse, 1998). Rounded back vowels following retroflex sibilants (Rau and Li, 1994) or alveolar sibilants (Jeng, 2006) tended to increase retroflex sibilant productions. Sociolinguistic studies (Chen, 1991; Luo, 1991; Rau and Li, 1994; Tse, 1998; Jeng, 2006) revealed that individual and contextual factors influenced the usage frequency of retroflex sibilants in Taiwan Mandarin. Women, youth, highly educated people, and those who used Taiwan Mandarin (not Taiwan Southern Min) at home tended to use prescriptive retroflex pronunciations. Formal speech styles and serious speech content also induced more frequent occurrences of prescriptive retroflex pronunciations. Ma (2006) experimentally investigated the perceptual awareness of attributing retroflexion to higher socioeconomic status among Taiwan Mandarin speakers. Chung (2006) also observed that full retroflexing was socially marked as compared to the intermediate forms which were the default covert prestige forms for all groups of speakers in daily conversations. Hypercorrection (i.e., the incorrect substitution, in a prescriptive sense, of retroflex sibilants for the corresponding alveolar sibilants) is not uncommon in Taiwan Mandarin, which indicates speakers' conscious association with interlocutors and social registers in the speech community (Kubler, 1985; Chung, 2006).

Taiwan Mandarin rimes contain only alveolar or velar nasal codas in the closed syllables (Chen, 1973). Even though both of nasal codas follow nearly all kinds of nuclei in common, some Taiwan Mandarin phonological rules change the surface realizations of the vowel-nasal combinations in the rimes. First, the low nucleus /a/ is raised to [ɛ] between a front vowel /i/ or /y/ and an alveolar nasal coda /n/. Second, the rounded high nucleus /u/ is lowered to [o] when preceded by a velar nasal coda /ŋ/. Third, the rounded high nucleus /y/ is diphthongized to [io] when preceded by a velar nasal coda /ŋ/. Fourth, the low nucleus /a/ is backed to [ɑ] when preceded by a velar nasal coda /ŋ/. In addition, there is a phonological gap of rimes as /*yaŋ/.

By observing that place distinction of nasal codas is frequently dropped in Taiwan Mandarin, researchers have shown intensive interests in determining the direction of nasal coda neutralization in Taiwan Mandarin, as other dialectal and historical linguists used to debate the same issue when they attempted to reconstruct nasalization processes in Chinese history (Chen, 1973; Zee, 1985; Hess, 1990). Some researchers argued for a single unidirectional merging direction [ŋ] > [n], regardless of the preceding vowels (Kubler, 1985; Tse, 1992; Chiou, 1997; Yang, 2007). Still others claimed that nasal codas tended to be alveolarized before a mid vowel [ə], but velarized before a high vowel [i] (Wu, 1985; Chen, 2000; Lin, 2002; Hsu and Tse, 2007; Lai, 2009). Among the studies which proposed bidirectionality of nasal coda neutralization, they even differed in whether alveolarization or velarization is the leading trend in Taiwan Mandarin.

Regarding the influencing factors of nasal coda neutralization in Taiwan Mandarin, researchers have investigated phonological, stylistic, and sociolinguistic variables. Phonologically speaking, Chiou (1997) found that the following coronal consonants enhanced alveolarization of nasal codas, even though the tone values and the prosodic boundaries barely influenced nasal coda neutralization. Tse (1992) experimentally investigated that nasal coda neutralization was more frequent when pronounced in the phrases or sentences than in the minimal-pair word lists. Sociolinguistic studies also observed that nasal coda neutralization in Taiwan Mandarin was conditioned by the age, gender, social status, education level, and ethnicity in an intertwining way (Kubler, 1985; Chen, 1991; Tse, 1992; Yueh, 1992). However, a latest study by Hsu and Tse (2007) found that those sociolinguistic variables previously claimed to influence nasal coda neutralization in Taiwan Mandarin have been leveled out today.

2. Effects of lexical frequency and neighborhood density

In a great deal of psycholinguistic studies, the factors related to the organization and activation of the words in the mental lexicon have been shown to influence speech perception and speech production. One factor is lexical frequency that counts how often words are used (e.g., *buy* and *goal* are more frequent than *bough* and *foal*). Common words are both recognized (Oldfield and Wingfield, 1965; Luce and Pisoni, 1998) AND produced (Geffen and Luszcz, 1983; Dell, 1990; Jescheniak and Levelt, 1994; Bonin and Fayol, 2002) faster than rare words. Rare words are more susceptible to speech errors than common words (Dell, 1990; Stemberger and McWhinney, 1986).

The other factor relevant to lexical frequency is neighborhood density that counts how many other words that are phonologically similar to the target words (e.g., *cat* and *lick* have more lexical neighbors than *quiz* and *purge*). Words from dense neighborhoods are recognized more slowly and less accurately than those from sparse neighborhoods (Goldinger, Luce, and Pisoni, 1989; Luce, 1986; Luce and Pisoni, 1998; Vitevitch and Luce, 1998, 1999). Nevertheless, neighborhood density has provided contradictory evidence in speech production. Some studies found that dense neighborhoods made speech production more accurate and faster than sparse neighborhoods (Gordon, 2002; Gordon and Dell, 2001; Harley and Brown, 1998), whereas other studies showed that spare neighborhood words were produced more quickly and with shorter durations than dense neighborhood words (Luce and Pisoni, 1998, Vitevitch and Luce, 1998).

Furthermore, lexical effects have been found to influence phonetic variation. While high lexical frequency words or phrases are more subject to reduction processes (Zipf, 1935; Balota, Boland, and Shields, 1989; Bybee and Hopper, 2001; Pierrehumbert, 2002; Myers and Li, 2009), dense neighborhoods cause speech production to move toward the canonical or even exaggerated forms as compared with sparse neighborhoods (Munson and Solomon, 2004; Wright, 2004; Munson, 2007).

Given the lexical influences in spoken word processing, our study attempts to explore how lexical factors affect two Taiwan Mandarin neutralization patterns during speech production in. Furthermore, the examination of lexical factors in neutralization also aims at explicating the cognitive mechanism underlying neutralization. One way to account for neutralization during speech production is based on the assumption that phonetic variability is the product of phonological processes transmitting continuous articulatory parameters into overt speech actualization (Bybee, 2001). Thus for frequent words and sparse neighborhood words, proper speech preparation can increase articulatory velocity which in turn can yield a larger amount of reduced realizations in speech production. Alternatively, phonetic variability can be the result of lexical selection that bypasses phonological planning processes; speakers can select an exemplar of a word from memory, and then use it, averaged with similar exemplars, as a goal for speech production (Pierrehumbert, 2002). Since frequent words and sparse neighborhood words contain more memories of reduced traces, as speakers have accumulated from prior speech experiences, those words naturally surface with more reduction. To investigate these hypotheses behind neutralization in Taiwan Mandarin, we thus conducted a production experiment as in the next section.

3. A production experiment

3.1 Participants

Twenty-four undergraduates from National Chung Cheng University participated in this experiment by receiving a reasonable fee. A half of them were women. In each gender, a half of them spoke Taiwan Mandarin as home while other half spoke Taiwan Southern Min at home. Ages of participants ranged from 19 to 23 (M=20.7, SD=1.5). All of participants were fluent Taiwan Mandarin speakers by self-evaluation.

3.2 Stimuli

The stimuli were 705 monosyllables with 308 pronounced with one of six sibilant onsets [ts, ts^h, s, tʂ, $tʂ^h$, ʂ] and 527 pronounced with one of two nasal codas [n, ŋ]. Some of the monosyllables contained both of a sibilant onset and a nasal coda. All of the monosyllables were presented in Chinese characters as the majority of 80 pretest participants chose the first instances among the homophones upon seeing the Romanization forms (zhuyin fufao). An equal amount of monosyllables not pronounced with those sibilant onsets and nasal codas were also prepared as the fillers.

3.3 Procedure

The experiment took place in a double-walled sound-attenuated room for approximately 30 minutes. The visual stimuli were presented on a 17-inch monitor. Participants were instructed to respond to a microphone on a stand at the moment when a response-triggering tone appeared. The experiment was preceded by a familiarization section of ten trials. Each participant had to read aloud the Chinese characters either concurrent with a response-triggering tone or 1000 ms before a response-triggering tone. The immediate-response and delayed-response conditions were counterbalanced by both of stimuli and participants. The experiment always advanced trials two seconds after the tone appeared. The experiment was processed by the DMDX program (Forster, 2009). The program also recorded production latency from the onset of Chinese characters to the initiation of participants' responses; the responses were simultaneously digitized at a sampling rate of 22 kHz with 16-bit quantization.

3.4 Data Preparation

The data were segmented and transcribed using Praat (Boersma and Weenink, 2009). The segmentation procedure was machine-made with an aid of Praat built-in segmentation function which tracked syllabic and segmental boundaries automatically. The boundaries were then readjusted manually by the author. For the sibilant onsets, the acoustic measurements comprised the information from the fricative noises (spectral moments, peak/slope parameters, duration, and average intensity) and the adjacent vowels (the first three formant frequencies on the onset of following adjacent vowels). For the nasal codas, the acoustic measurements comprised the information from the nasal consonants (amplitude differences of the nasal formant and the first formant, nasal duration, and average nasal intensity) and adjacent vowels (the first three formants at the offsets of the preceding nucleus vowels).

4. Statistical analyses

In order to maximize the power of our analysis, we entered all the acoustic measurements (after orthogonalized by Principal Component Analysis) into a linear mixed-effects model (LME; Baayen, 2008). The LME model also enhances statistical sensitivity by allowing us to include both of random variables (participants and items) among the other fixed variables in the model. The current likelihood ratio test revealed that by-participant-and-item model always fitted the dataset better than by-participant-only model or by-item-only model. The tool that we used for modeling the LME analysis was the *lmer* function in the *lme4* package (Bates, Maechler, and Dai, 2008) in R (R Development Core Team, 2012). The *p* values from the LME model were simulated by means of Markov

Chain Monte Carlo (MCMC) sampling 10,000 times using the *pvals.fnc* function in the *languageR* package (Baayen, 2008).

4.1 Predictors

In this subsection, we report the predictor variables that may influence Taiwan Mandarin neutralization.

4.1.1 Lexical Variables: Lexical Frequency and Neighborhood Density

As the most important predictor variables we were concerned with, we calculated lexical frequency and neighborhood density of all the monosyllabic stimuli. Lexical frequency was the calculation of the occurrences from two Chinese spoken corpora: *SUBTLEX-CH*: *Chinese Word and Character Frequencies Based on Film Subtitles* (Cai and Brysbaert, 2010) and *Taiwanese Putonghua Speech and Transcripts in the Linguistic Data Consortium* (Duanmu, Wakefield, Hsu, Qui, and Cristina, 1998). For a particular Chinese character, lexical frequency was the sum of occurrences from both of spoken corpora. As following the single segment edit distance metric (Luce, 1986; Luce and Pisoni, 1998), a phonological neighbor was defined as any syllable derived by substituting, deleting, or inserting a single segment or tone. Given the spoken corpora yielded the frequencies of all homophonic Chinese characters, neighborhood density of a monosyllable was the sum of occurrences of all the neighboring homophones from both of the spoken corpora. Lexical frequency (**logFreqSpoken**) and neighborhood density (**logFreqSpoken**) were logarithm normalized before being input to statistical analyses.

4.1.2 Orthographic Variables: Familiarity of Characters and Zhuyin Fuhao

As in previous studies, orthography tended to preserve phonetic differences in the neutralization process of two segments (Fourakis and Iverson, 1984; Jassem and Richter, 1989; Warner, Jongman, Sereno, and Kemps, 2004; Warner, Good, Jongman, and Sereno, 2006). Apart from that, the orthographic variables hereby functioned factoring out the prelexical visual identification process, once it confounded with spoken word processes, during the present production task. We assessd orthographic influences by conducting two sets of thermometer judgment tasks to obtain the familiarities of Chinese characters and Romanization forms (zhuyin fuhao). Both familiarity scores for Chinese characters (**logFamChar**) and zhuyin fuhao (**logFamZhu**) were logarithm normalized before being input to statistical analyses.

4.1.3 Competence-Based Variables: Accuracy of Spelling out Monosyllables

An assumption related to pronunciation variability from prior sociolinguistic Taiwan Mandarin studies is that speakers might produce canonical or deviant forms biased from their prescriptive knowledge about the pronunciation of an orthographic form. Failure to produce the canonical forms might therefore be due the lack of prescriptive knowledge of the canonical forms. For instance, Taiwan Mandarin speakers might mispronounce retroflex or alveolar sibilants simply because they never took note of the prescriptive pronunciations of retroflex or alveolar sibilants during the period that they learned Chinese characters. Accordingly, we assessed the prescriptive knowledge by asking participants to spell out the pronunciations of the Chinese character stimuli using the Romanization forms (zhuyin fuhao). Accuracy ratio (**charMatchZhu**) was averaged across participants, i.e., the proportion of whether participants spelled out the target sibilants or nasals of the stimuli accurately.

4.1.4 Contextual Variables: Features of Segments and Adjacent Vowels

Previous Taiwan Mandarin studies have observed that the occurrences of segmental neutralization coordinated with neighboring segmental features. Acoustic properties of retroflex and alveolar sibilants in Taiwan Mandarin varied with their co-present aspiration and affrication (Tse, 1998; Jeng, 2006). Previous studies in Taiwan Mandarin also showed that lowness or highness of the preceding vowels influenced nasal coda neutralization to differential extents (Kubler, 1985; Wu, 1985; Chen, 1991; Tse, 1992; Chiou, 1997; Lin, 2002; Hsu and Tse, 2007). We thus took the contextual variables into account to prevent these variables from confounding with the lexical variables in affecting neutralization. For coding the contextual features of sibilant onsets, we specified the features of **retroflexion**, **aspiration** and **affrication** of the sibilant onsets and the highness of the following vowels (**vHigh**). For coding the contextual features of the nasal codas, we specified the **alveolar** places of articulation of the nasal codas and the lowness of the preceding vowels (**vLow**).

4.1.5 Processing Variables: Production Latency and Response Duration

One of the present goals in our study is to examine whether lexical influences in neutralization come from real-time planning processes or lexically selective processes. Previous studies have shown that lexical frequency affected access of lexical items (Dell, 1990; Jescheniak and Levelt, 1994; Bonin and Fayol, 2002) and the lexical effects could be extended to postlexical pronunciation processes (Balota and Chumbley, 1985; Goldinger, Azuma, Abramson, and Jain, 1997). Some other studies also have found that high lexical frequency reduced response duration (Wright, 1979; Geffen and Luszcz, 1983, Kawamoto, Kello, Higareda, and Vu, 1999; Jurafsky, Bell, Gregory, and Raymond, 2001; Munson and Solomon, 2004). Moreoever, neutralization can be the consequence of duration-dependent undershoot or overshoot; for instance, Moon and Lindblom (1994) found that shorter vowels tended to be produced closer to the Euclidian center of the F1/F2 space than longer vowels. Such being the case, neutralization can be the accumulation of a sequential temporal transition from lexical access, phonological processes, to articulation during speech production. Accordingly, we specified two processing variables: production latency or reaction times (**RT**) and syllable duration (**sylDur**) that roughly indicated the processes of speech preparation and speech actualization, respectively.

4.2 Dependent Measurements

To deal with multidimensional acoustic measurements (3.4), statistical technique of Principal Components Analysis (PCA) maximizes the explanation of the variances while shrinking down the dimensions of the variances in the dataset to very few crucial ones (Baayen, 2008). The output of PCA is a transformed matrix with the reduced number of uncorrelated principal components. In practice, we only selected the first principal components (PC1) which accounted for 81% of the variances of the sibilant measurements or 70% of the variances of the nasal measurements, respectively. Pearson's product-moment correlation tests showed that all of the acoustic measurements significantly correlated with the first principal components, proving that the first principal components were realizably representative of all of the acoustic measurements.

4.3 Results

The production experiment on twenty-four subjects yielded a total of 7,392 sibilant tokens (308 words × 24 subjects) and 12,648 nasal tokens (308 words × 24 subjects). Prior to analysis, 330 (4.5%) sibilant errors and 386 (1.0%) nasal errors were discarded due to mispronunciations and non-responses in the trials. As initializing the analysis of the dataset, we performed linear mixed-effects (LME) models for the immediate-response condition and the delayed-response condition separately.

4.3.1 Sibilant Onset Neutralization

In the immediate-response condition where speakers read the stimuli simultaneously with the response cues, the main effects in the LME model (1) indicate that the contextual (retroflexion, aspiration, and vocalic highness), orthographic (familiarity of zhuyin fuhao), and lexical variables (neighborhood density) influenced the way that speakers pronounced sibilants. Retroflexion, aspiration, vocalic highness, and higher neighborhood density increased retroflex productions, while higher familiarity of zhuyin fuhao increased alveolar productions.

To assess the neutralization of sibilant onsets, we cared more about the interactions of retroflexion with the other predictors in the LME model. For the contextual variables, affrication, aspiration, and vocalic highness enhanced neutralization by increasing retroflex productions. For the orthographic variable, higher familiarity of zhuyin fuhao raised neutralization by increasing retroflex productions. For the processing variables, faster reaction times caused more neutralization by increasing alveolar productions. No interactions of retroflexion with the other predictors, including those with lexical variables, were found. The partial effects of the lexical variables and the processing variables of two sibilant onsets (with 95% confidence intervals) are illustrated in (3).

(1) Linear mixed-effects regression model of the predictors for PC1 on sibilant onsets in the immediate-response condition.

N = 3,580	*Estimate*	*Std. Error*	*t*	*p* (MCMC)
(Intercept)	3270.844	1001.977	3.264	.0011*
retroflexion	-2552.969	1216.873	-2.098	.0360*
aspiration	-724.143	135.099	-5.360	<.0001*
affrication	-214.078	149.037	-1.436	.1510
vowelHigh	-900.795	112.082	-8.037	<.0001*
logFamChar	-1040.302	580.659	-1.792	.0733
logFamZhu	921.863	407.140	2.264	.0236*
charMatchZhu	-627.475	537.757	-1.167	.2434
RT	55.846	211.218	.264	.7915
sylDur	116.548	460.869	.253	.5131
logFreqSpoken	-60.480	79.635	-.759	.4476
logNeighborSpoken	-452.595	147.447	-3.070	.0022*
retroflex:aspiration	386.662	181.510	2.130	.0332*
retroflex:affrication	328.911	193.611	1.699	.0894
retroflex:vowelHigh	787.099	152.001	5.178	<.0001*
retroflex:logFamChar	1642.838	841.622	1.952	.0510
retroflex:logFamZhu	-1166.230	541.713	-2.153	.0314
retroflex:charMatchZhu	546.838	711.415	.769	.4421
retroflex:RT	587.293	224.423	2.617	.0089*
retroflex:sylDur	576.691	519.161	1.111	.2667
retroflex:logFreqSpoken	66.338	101.419	.654	.5130
retroflex:logNeighborSpoken	269.856	182.930	1.475	.1400

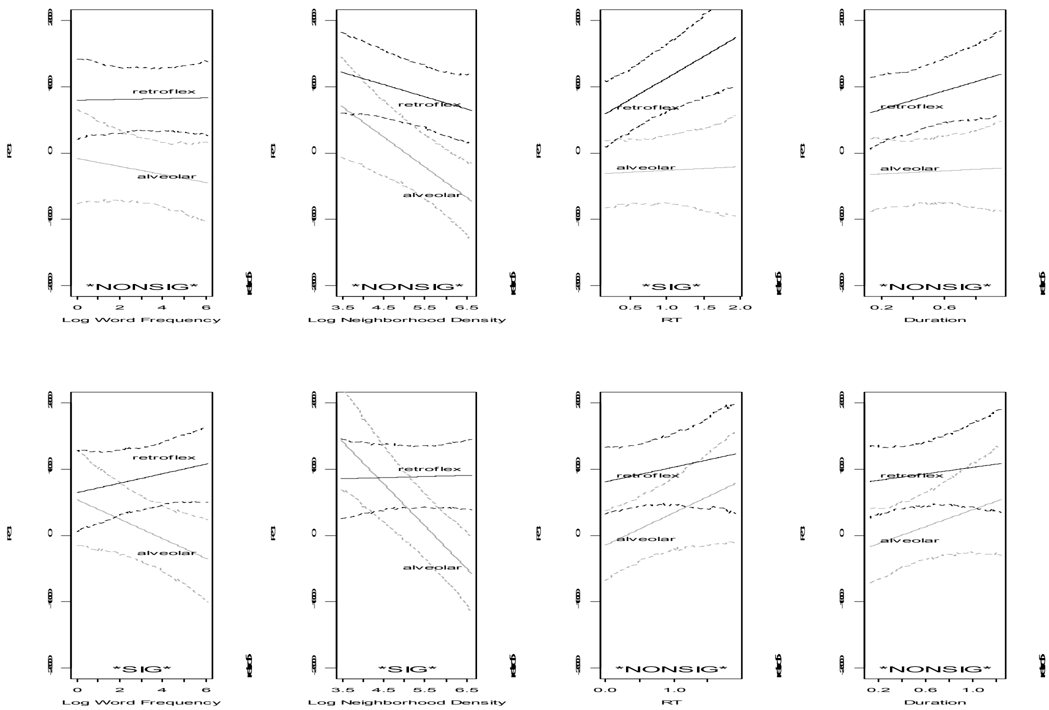
In the delayed-response condition where speakers read the stimuli 1000 ms before the response cues, the main effects in the LME model (2) indicate that aspiration, vocalic highness and higher neighborhood density increased retroflex productions, while higher reaction times increases alveolar productions.

To investigate the neutralization of sibilant onsets, we further pay attention to the interactions of retroflexion with the other predictors in the LME model. It was found that vocalic highness promoted neutralization by increasing retroflex productions; lower lexical frequency and lower neighborhood density enhanced neutralization by either drawing two sibilant productions closer to the intermediate place of articulation or increasing retroflex productions. No interactions of retroflexion with the other predictors, including those with the processing variables, were found.

(2) Linear mixed-effects regression model of the predictors for PC1 on sibilant onsets in the delayed-response condition.

N = 3,482	*Estimate*	*Std. Error*	*t*	*p* (MCMC)
(Intercept)	3675.528	1028.004	3.575	.0004*
retroflexion	-2141.342	1260.294	-1.699	.0894
aspiration	-691.939	139.564	-4.958	<.0001*
affrication	-267.813	152.313	-1.758	.0788
vowelHigh	-1212.547	116.661	-10.394	<.0001*
logFamChar	-404.249	604.561	-.669	.5038
logFamZhu	731.679	425.119	1.721	.0853
charMatchZhu	-473.301	556.543	-.850	.3951
RT	488.371	240.589	2.030	.0424*
sylDur	640.689	446.787	1.434	.1517
logFreqSpoken	-147.514	82.871	-1.780	.0752
logNeighborSpoken	-147.514	1260.294	-4.119	<.0001*
retroflex:aspiration	290.663	188.437	1.542	.1230
retroflex:affrication	373.439	199.685	1.870	.0615
retroflex:vowelHigh	1158.213	158.144	7.324	<.0001*
retroflex:logFamChar	-705.450	875.063	-.806	.4202
retroflex:logFamZhu	-899.734	566.769	-1.587	.1125
retroflex:charMatchZhu	1283.677	740.876	1.733	.0832
retroflex:RT	-268.211	242.727	-1.105	.2692
retroflex:sylDur	-395.124	499.334	-.791	.4288
retroflex:logFreqSpoken	219.038	105.365	2.079	.0377*
retroflex:logNeighborSpoken	650.389	190.798	3.409	.0007*

(3) Partial effects of the lexical variables and processing variables on two sibilant onsets in the immediate- (upper) and delayed- (lower) conditions. An interaction of retroflexion with the other variable is marked *SIG* in the bottom of the graph.



4.3.2 Nasal Coda Neutralization

The results from the immediate-response condition (4) show no main effects in general nasal coda productions. Nevertheless, we found a number of interactions of alveolar (place of articulation) with the other predictors in the LME model. Non-low vowels enhanced neutralization by increasing alveolar productions. The processing variables i.e., slower reactions times and shorter response duration enhanced neutralization by increasing alveolar productions. The lexical variables i.e., higher lexical frequency and higher neighborhood density raised neutralization by drawing two sibilant productions closer to the immediate place of articulation. The partial effects of the lexical variables and the processing variables of two nasal codas (with 95% confidence intervals) are illustrated in (6).

(4) Linear mixed-effects regression model of the predictors for PC1 on nasal codas in the immediate-response condition.

N = 6,190	*Estimate*	*Std. Error*	*t*	*p* (MCMC)
(Intercept)	86.784	350.465	.248	.7804
alveolar	-358.572	540.859	-.663	.5074
vowelLow	46.052	23.695	1.944	.0520
logFamChar	120.178	126.120	.953	.3407
logFamZhu	-36.811	168.936	-.218	.8275
charMatchZhu	78.445	189.835	.413	.6795

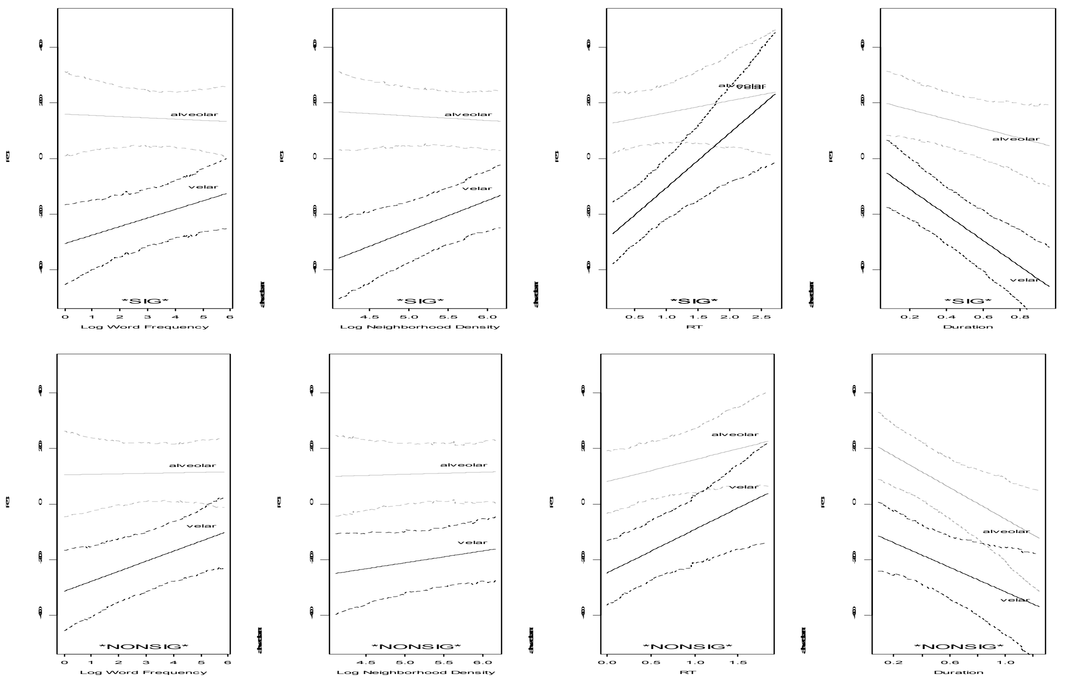
RT	43.350	48.838	.888	.3748
sylDur	-172.259	99.699	-1.728	.0841
logFreqSpoken	-4.484	17.951	-.663	.8027
logNeighborSpoken	-16.662	38.976	-.427	.6690
alveolar:vowelLow	86.439	35.863	2.410	.0160*
alveolar:logFamChar	-130.621	179.703	-.727	.4673
alveolar:logFamZhu	-287.167	268.304	-1.070	.284
alveolar:charMatchZhu	-155.273	333.413	-.466	.6414
alveolar:RT	153.075	55.061	2.780	.0055*
alveolar:sylDur	-292.660	122.414	-2.391	.0168*
alveolar:logFreqSpoken	135.248	24.370	2.446	.0048*
alveolar:logNeighborSpoken	127.061	56.116	2.264	.0236*

The results from the delayed-response condition (5) show the main effects of the contextual variable i.e., vocalic lowness and the processing variable i.e., response duration, indicating that vocalic lowness and longer response duration induced more velar productions. Crucially to our analysis, no interaction of alveolar (place of articulation) with the other predictors, including the lexical variables and the processing variables was found in the LME model.

(5) Linear mixed-effects regression model of the predictors for PC1 on nasal codas in the delayed-response condition.

N = 6,071	*Estimate*	*Std. Error*	*t*	*p* (MCMC)
(Intercept)	-40.676	338.431	-.120	.9043
alveolar	31.903	525.416	.061	.9516
vowelLow	68.909	23.139	2.978	.0029*
logFamChar	-66.050	123.003	-.537	.5913
logFamZhu	160.893	165.127	.974	.3299
charMatchZhu	54.227	184.033	.295	.7683
RT	78.703	50.472	1.559	.1190
sylDur	-284.552	95.243	-2.988	.0028*
logFreqSpoken	1.871	17.490	.107	.9148
logNeighborSpoken	7.977	38.034	.210	.8339
alveolar:vowelLow	16.211	34.914	.464	.6424
alveolar:logFamChar	112.244	174.211	.644	.5194
alveolar:logFamZhu	-495.013	261.706	-1.891	.0586
alveolar:charMatchZhu	-146.411	324.715	-.451	.6521
alveolar:RT	76.951	56.543	1.361	.1736
alveolar:sylDur	63.200	120.182	.526	.5990
alveolar:logFreqSpoken	34.308	23.715	1.447	.1480
alveolar:logNeighborSpoken	35.019	54.521	.642	.5207

(6) Partial effects of the lexical variables and processing variables on two sibilant onsets in the immediate- (upper) and delayed- (lower) conditions. An interaction of alveolar with the other variable is marked *SIG* in the bottom of the graph.



5. Discussion

This section discusses the findings from the production experiment. First, for retroflex sibilant neutralization, most of predictors (including the three contextual variables and neighborhood density) approximated neutralization to retroflex productions, while lexical frequency made neutralization close to the intermediate place of articulation between two sibilant onsets, but reaction times brought about neutralization for alveolar productions. For nasal coda neutralization, vocalic lowness, reaction times, and response duration activated alveolar productions during neutralization, while lexical frequency and neighborhood density triggered neutralization close to the intermediate place between two nasal codas. The present patterns were not inconsistent with the previous Taiwan Mandarin studies. Even though neutralization was usually expected to converge on the unmarked alveolar (vs. retroflex or velar) places of articulation, retroflex sibilant neutralization exhibited hypercorrection, i.e., the substitution of retroflex sibilants for alveolar sibilants, probably due to sociolinguistic awareness of Taiwan Mandarin speakers.

Second, higher lexical frequency and higher neighborhood density were correlated with higher nasal coda neutralization, whereas lower lexical frequency and lower neighborhood density were correlated with higher sibilant onset neutralization. The contrast was here found across the neutralization patterns, which differed from previous research in that higher lexical frequency enhanced reduction processes while higher neighborhood density induced hyperarticulated speech (e.g., Bybee and Hopper, 2001; Munson and Solomon, 2004; Wright, 2004; Munson, 2007). One of the explanations for this is that speakers' sociolinguistic awareness made sibilant onset neutralization likely to occur in infrequent words and dense neighborhood words since speakers tended to adopt hypercorrect retroflex productions in facing the uncertainty and ambiguity of the prescriptive pronunciations in those words. The other explanation is based on the finding that speakers were rather insensitive to response duration and reaction times in sibilant onset neutralization. In both of immediate- and delayed- response conditions, speech actualization as measured by response duration did not affect neutralization at all. Speech preparation as measured by reaction times in the immediate-response condition affected neutralization but in a contrary direction to lexical effects in neutralization processes. This plausibly implied that lexical effects in sibilant onset neutralization were neutral to or different from the prediction of lexical effects on ordinary reduction processes in previous studies. By contrast, nasal coda neutralization was compatible with the account of the real-time processing mechanism. Speakers spent more time in accessing neutralized words (consistent with lexical effects in yielding more alveolar productions) probably due to self-monitoring of lexical ambiguity during speech preparation. As speakers began to initialize speech implementation, neutralization increased with shorter response duration, suggesting that speakers indeed condensed nasal coda productions in facing time restriction.

Third, lexical effects in sibilant onset neutralization occurred only in the delayed-response condition, while lexical effects in nasal coda neutralization occurred only in the immediate-response condition. The common findings across two neutralization patterns were no processing effects (reaction times and response duration) in neutralization in the delayed-response condition where speakers were under less time restriction to set up speech production. Lexical effects in sibilant onset neutralization were thus a case in which speakers did not utilize real-time processing to realize neutralization since the lexical effects only occurred in the delayed-response condition; instead, lexical effects in sibilant onset neutralization were likely a selective process of realized forms directly from the mental lexicon. By contrast, nasal coda neutralization was lexically processed only in the immediate-response condition; as temporal pressure decreased in the delayed-response condition, lexical effects in neutralization, which required an intermediate time-driven phase to realize neutralization, thus disappeared.

Overall, the contrast between two neutralization patterns is expected if we assume that two neutralization patterns are subject to different processing mechanisms. Nasal coda neutralization is gradually constructed via the transition of articulatory processes by time. Sibilant onset neutralization is a consciously selective process during speech preparation in that speakers are aware of the relationship of lexical idiosyncrasy and hypercorrect retroflex productions. The lexical effects in sibilant onset neutralization are not incompatible with previous diachronic changes that began with infrequent words for a socio-cognitive reason (Bybee, 2000; Phillips, 2006). The present findings encourage us to explore more diverse processing mechanisms of phonetic patterns even if they might look not different at first glance.

References

Baayen, R. H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistic Using R*. Cambridge: Cambridge University Press.

Balota, D. A., J. E. Boland, and L. W. Shields. 1989. Priming in Pronunciation: Beyond Pattern Recognition and Onset Latency. *Journal of Memory and Language* 28: 14-36.

Balota, D. A., and J. I. Chumbley. 1984. Are Lexical Decisions A Good Measure of Lexical Access? The Role of Word frequency in the Neglected Decision Stage. *Journal of Experimental Psychology: Human Perception and Performance* 10: 340-357.

Bates, D. and M. Maechler, and B. Dai. 2008. Lme4: Linear Mixed-Effects Model Using S4 Class. R package.

Boersma, P., and D. Weenink. 2009. Praat: Doing Phonetics by Computer (Version 5.1.04). Retrieved from http://www.praat.org/.

Bonin, P., and M. Fayol. (2002). Frequency Effects in the Written and Spoken Production of Homophonic Picture Names. *European Journal of Cognitive Psychology* 14(3): 289-313.

Bybee, J. 2000. The Phonology of the Lexicon: Evidence from Lexical Diffusion. In M. Barlow, and S. Kemmer, eds. *Usage-based Models of Language*, 65-86, Stanford: CSLI.

Bybee, J. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.

Bybee, J., and P., Hopper, eds. 2001. *Frequency and the Emergence of Linguistic Structure*. Amersterdam: John Benjamins.

Cai, Q., and M. Brysbaert. 2010. SUBTLEX-CH: Chinese Word and Character Frequencies Based on Film Subtitles. *PLoS ONE* 5(6): 1-8.

Chen, C.-Y. 1991. *Shendiao de zhuanbian yu kuosan: Taipei butong nianlingqun de quyang* [聲調的轉變與擴散: 台北不同年齡群的取樣] [The tonal change and diffusion: From the sampling of different age groups in Taipei]. *Journal of Chinese Language Teacher Association* 16(1): 69-99.

Chen, M. 1973. Cross-Dialectal Comparison: A Case Study and Some Theoretical Considerations. *Journal of Chinese Linguistics* 1(1): 38-63.

Chen, M. Y. 2000. Acoustic Analysis of Simple Vowels Preceding A Nasal in Standard Chinese. *Journal of Phonetics* 28: 43-67.

Cheng, C. C. 1973. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.

Chiou, F.-D. 1997. The Phonetic Realization of Final Engma in Taipei. *U. Penn Working Papers in Linguistics* 4(1): 203-227.

Chung, K. S. 2006. Hypercorrection in Taiwan Mandarin. *Journal of Asian Pacific Communication* 16(2): 197-214.

Dell, G. S. 1990. Effects of Frequency and Vocabulary Type on Phonological Speech Errors. *Language and Cognitive Processes* 5: 313-349.

Duanmu, S., G. H. Wakefield, Y.-P. Hsu, S.-P. Qui, and G. R. Cristina. 1998. *Taiwanese Putonghua Speech and Transcripts*. *Linguistic Data Consortium*. http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC98S72.

Forster, J. 2009. *DMDX Display Software* (Version 3.3.11). Retrieved from http://www.u.arizona.edu/~kforster/dmdx/download.htm.

Fourakis, M., and K. Iverson. 1984. On the 'Incomplete Neutralization' of German final obstruents. *Phonetica* 41: 140-149.

Geffen, G., and M. A. Luszcz. 1983. Are Spoken Durations of Rare Words Longer than Those of Common words? *Memory and Cognition* 11(1): 13-15.

Gordon, J. K., and G. S. Dell. 2001. Phonological Neighborhood Effects: Evidence from Aphasia and Connectionist Modeling. *Brain and Language* 79: 21-23.

Goldinger, S. D., T. Azuma, M. Abramson, and P. Jain. 1997. Open Wide and Say "Blah": Attentional Dynamics of Delayed Naming. *Journal of Memory and Language* 37: 190-216.

Goldinger, S. D., P. A. Luce, and D. B. Pisoni. 1989. Priming Lexical Neighbors of Spoken Words: Effects of Competition and Inhibition. *Journal of Memory and Language* 28: 501-518.

Gordon, J. K. 2002. Phonological Neighborhood Effects in Aphasic Speech Errors: Spontaneous and Structured Contexts. *Brain and Language* 82: 113-145.

Harley, T. A., and H. E. Brown. 1998. What Causes A Tip-of-the-Tongue State? Evidence for Lexical Neighborhood Effects in Speech Production. *British Journal of Psychology* 89: 151-174.

Her, O.-S. 2009. Language and Group Identity: On Taiwan Mainlanders' Mother Tongue and Taiwan Mandarin. *Language and Linguistics* 10(2): 375-419.

Hess, S. 1990. Universals of Nasalization: Development of Nasal Finals in Wenling. *Journal of Chinese Linguistics* 18(1): 44-94.

Hsu, H.-J., and J. K.-P. Tse. 2007. Syllable-Final Nasal Mergers in Taiwan Mandarin – Leveled but Puzzling. *Concentric: Studies in Linguistics* 33(1): 1-18.

Jassem, W., and L. Richter. 1989. Neutralization of Voicing in Polish Obstruents. *Journal of Phonetics* 17, 317-325.

Jeng, J.-Y. 2006. The Acoustic Spectral Characteristics of Retroflex Fricatives and Affricates in Taiwan Mandarin. *Journal of Humanistic Studies* 40(1): 27-48.

Jescheniak, J. D., and J. M. Levelt. 1994. Word Frequency Effects in Speech Production: Retrieval of Syntactic Information and of Phonological Form. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20: 824-843.

Jurafsky, D., A. Bell, M. Gregory, and W. D. Raymond. 2001. Probabilistic Relations between Words: Evidence from Reduction in Lexical Production. In J. Bybee, and P. Hopper, eds., *Frequency and the Emergence of Linguistic Structure*, 229-254, Amsterdam: John Benjamin.

Kawamoto, A. H., C. T. Kello, I. Higareda, and V. Q. Vu. 1999. Parallel Processing and Initial Phoneme Criterion in Naming Words: Evidence from Frequency Effects in Onsets and Rime Duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25(2): 362-382.

Kubler, C. C. 1985. The Influence of Southern Min on the Mandarin of Taiwan. *Anthropological Linguistics* 27(2): 156-176.

Lai, Y.-H. 2009. Acoustic Correlates of Mandarin Nasal Codas and Their Contribution to Perceptual Saliency. *Concentric: Studies in Linguistics* 35(2): 143-166.

Lin, C. C. 2002. Nasal Endings of Taiwan Mandarin: Production, Perception, and Linguistic Change. Paper presented at the 35th International Conference on Sino-Tibetan Language and Linguistics, Arizona State University, Tempe.

Luce, P. A. 1986. Neighborhoods of Words in the Mental Lexicon. *Research on Speech Perception, Technical Report No. 6*. Speech Research Laboratory, Indian University, Bloomington, IN.

Luce, P. A., and Pisoni, D. B. 1998. Recognizing Spoken Words: The Neighborhood Activation Model. *Ear and Hearing* 19(1): 1-36.

Ma, P.-Y. 2006. MA. Thesis, National Chung Cheng University, Chiayi, Taiwan, *Social Stereotypes and Accent Perception in Taiwan*. Unpublished.

Moon, S.-J., and B. Lindblom. 1994. Interaction between Duration, Context, and Speaking style. *Journal of the Acoustical Society of America* 96(1): 40-55.

Munson, B. 2007. Lexical Access, Lexical Representation, and Vowel Production. In J. Cole, and J. I. Hualde, eds., *Laboratory phonology 9*, 201-227. Berlin: Mouton de Gruyter.

Munson, B., and N. P. Solomon. 2004. The Influence of Phonological Neighborhood Density on Vowel Articulation. *Journal of Speech, Language, and Hearing Research* 47: 1048-1058.

Myers, J., and Y. Li. 2009. Lexical Frequency Effects in Taiwan Southern Min Syllable Contraction. *Journal of Phonetics 37*: 212-230.

Oldfield, R. C., and A. Wingfield. 1965. Response Latencies in Naming Objects. *Quarterly Journal of Experimental Psychology* 17: 273-281.

Phillips, B. S. 2006. Word Frequency and Lexical Diffusion. New York: Palgrave MacMillan.

Pierrehumbert, J. B. 2002. Word-Specific Phonetics. In C. Gussenhoven, and N. Warner, eds., *Laboratory phonology 7*, 101-139. Berlin: Mouton de Gruyter.

R Development Core Team. 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN: 3-900051-17-0, URL <http://www.R-project.org>.

Rau, D.-H., and M.-C. Li. 1994. Phonological Variation of /tʂ/, /tʂʰ/, /ʂ/ in Mandarin Chinese. *Proceedings of the Fourth International Conference on Chinese Language Pedagogy, Chientan, Taipei*, 346-366.

Stemberger, J. P., and B. MacWhinney. 1986. Frequency and the Lexical Storage of Regularly Inflected Forms. *Memory and Cognition*, *14*, 17-26.

Tsao, F.-F. 2000. *Taishi Riyu yu Taiwan Guoyu – yibai nian lai zhai Taiwan fashen de liang ke yuyian Jiechu shili* [台式日語與台灣國語 – 一百年來在台灣發生的兩個語言接觸實例] [Taiwanized Japanese and Taiwan Mandarin – Two cases studies of language contact during the past hundred years in Taiwan]. *Chinese Studies* 18: 273-297.

Tse, J. K.-P. 1992. Production and Perception of Syllable Final [n] and [ŋ] in Mandarin Chinese: An Experimental Study. *Studies in English Literature and Linguistics* 18: 143-156.

Tse, J. K.-P. 1998. *Taiwan diqu nianqingren ㄓㄔㄕ yu ㄗㄘㄙ zhende bu fen ma?* [台灣地區年輕人ㄓㄔㄕ與ㄗㄘㄙ真的不分嗎?] [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.

Tung, C.-S. 1994. *Taiwan Hanyu fangye yingxiang xia de ruogan shengmu bianti chugao* [台灣漢語方言影響下的若干『國語』聲母變體初稿] [A preliminary study on some variants of Taiwan Chinese dialect influenced "Guoyu" initials]. *Proceedings of the Fourth International Conference on Chinese Language Pedagogy, Chietan, Taipei*, 323-343.

Vitevitch, M. S., and P. A. Luce. 1998. When Words Compete: Levels of Processing in Perception of Spoken Words. *Psychological Science* 9: 325-329.

Vitevitch, M. S., and P. A. Luce. 1999. Probabilistic Phonotactics and Neighborhood Activation in Spoken Word Recognition. *Journal of Memory and Language 40*: 374-408.

Warner, N., E. Good, A. Jongman, and J. Sereno. 2006. Orthographic vs. Morphological Incomplete Neutralization Effects. *Journal of Phonetics* 34: 285-293.

Warner, N., A. Jongman, J. Sereno, and R. Kemps. 2004. Incomplete Neutralization and Other Sub-Phonemic Durational Differences in Production and Perception: Evidence from Dutch. *Journal of Phonetics* 32: 251-276.

Wright, C. E. 1979. Duration Differences between Rare and Common Words and Their Implications for the Interpretation of Word Frequency Effects. *Memory and Cognition* 7(6): 411-419.

Wright, R. 2004. Factors of Lexical Competition in Vowel Articulation. In J. Local, R. Ogden, and R. Temple, eds., *Phonetic Intpretation: Papers in Laboratory VI*, 75-87. Cambridge: Cambridge University press.

Wu, K.-H. 1985. *Guoyu huayin cai Taiwan: Muqian qunshi yu iban chuou zhi tangtao* [國語發音在台灣:目前趨勢與一般錯誤之探討] [The pronucnciation of Standard Chinese in Taiwan: The invesigation of the current trend and the common errors]. *Proceedings of the First International Conference on the Teaching of Chinese as A Second Language*, 413-425. Taipei: Word Chinese Language Association.

Yang, J.-H. 2007. The Role of Sound Change in the Speech Recognition System: A Phonetic Analysis of the Final Nasal Shift in Mandarin. Paper presented at the Conference on Computational Linguistics and Speech Processing, Natioanl Taiwan University, Taipei.

Yueh, L.-C. 1992. MA. Thesis, Fu Jen Catholic University, Taipei, Taiwan, *The Drift of the Velar Nasal Ending in Taiwan Mandarin: A Sociolinguistic Survey*. Unpublished.

Zee, E. 1985. Sound Change in Syllable Final Nasal Consonants in Chinese. *Journal of Chinese Linguistics* 13(2): 290-330.

Zipf, G. K. 1935. *The Psycho-Bbiology of Language*. Boston: Houghton Mifflin.

Ying-Shing Li
Institute of Linguistics, Academia Sinica
128 Academia Road, Section 2, Nangang
Taipei 115, Taiwan

yingshing_li@yahoo.com