Concentric: Studies in Linguistics
43.1 (May 2017): 1-23 DOI: 10.6241/concentric.ling.43.1.01

The Influence of Dialect Information on the Perception of the Mandarin Alveolar-Retroflex Contrast*

Yung-Hsiang Shawn Chang
National Taipei University of Technology

Socioperceptual research has shown that speech perception can be biased by the perceived dialect of a speaker. While these studies have focused on vowels, the current study investigates whether fricative perception exhibits similar dialect effects. The Mandarin alveolar-retroflex contrast, often used for distinguishing between Beijing and Taiwan Mandarin, was the manipulated feature. When presented with a Beijing or Taiwan label, Beijing and Taiwanese listeners identified various tokens along the /ʂa-sa/ and /ʂu-su/ continua. The results showed no evidence that the dialect labels affected listeners' responses. The roles of dialect stereotypes and attitudes were also examined, but none of these variables influenced the identification results. The absence of dialect labeling effects in fricative perception may be due to the lack of control of the participants' alveolar and retroflex production, and to the limited dialect-distinguishing cues contained in Mandarin fricatives alone. In light of other research which also reported limited evidence for social information on perception, it is suggested that future research should investigate under what circumstances (e.g., control of listener variables, choice of stimuli) such dialect labeling effects in consonant perception might be expected.

Key words: social information, perceptual bias, Mandarin alveolar-retroflex contrast, dialect stereotype

1. Introduction

Besides linguistic information, speech signals also contain social-indexical information such as the speaker's dialectal background. Both sources of information are encoded in memory and interact with each other in perception: phonological variation in speech affects listeners' dialect characterization decisions (e.g., van Bezooijen & Gooskens 1999, Purnell, Idsardi & Baugh 1999, Clopper & Pisoni 2004, 2007); listeners adapt their perception as a function of the regional accent of the precursor information (e.g., Evans & Iverson 2004, Maye, Aslin & Tanenhaus 2008). Social-indexical information can further interfere with linguistic processing in that a listener's beliefs or stereotypes about the speaker's dialect have a biasing effect on their speech perception, a phenomenon that Niedzielski (1999) has demonstrated with perceived Canadian raising. Niedzielski asked her Detroit listeners to match the /aʊ/ contained in words produced by a Detroit speaker to a continuum of resynthesized

* I am grateful to two anonymous reviewers for their valuable comments on the earlier version of this study. I also thank Dr. Rixin Zhao at Beijing Language and Culture University for Beijing participant recruitment. This research was supported by the Ministry of Science and Technology, Taiwan, R.O.C. under Grant No. NSC 103-2410-H-027-003.

vowel tokens (varying F1 and F2). Before the task, half the listeners were told that the speaker was from Detroit, and the other half were told that the speaker was from Canada. It was found that the listeners given the Canadian label chose raised-diphthong tokens, whereas the listeners in the Detroit-label group consistently selected the unshifted vowels as the best match, even though both groups heard the same raised variants produced by the Detroit speaker. The results indicate that if listeners believe a speaker is from a certain dialect region, they may expect the speaker to sound a certain way and therefore are more likely to hear the speaker that way. Following Niedzielski's (1999) experimental paradigm, Hay, Nolan & Drager (2006) reported similar biasing effects of regional labels (Australian vs. New Zealander) on the perception of the /ɪ/ vowel. In addition to explicit regional descriptions, Hay & Drager (2010) further demonstrated that objects conveying the speaker's dialect in the listener's environment could also affect how a vowel is perceived. They covertly exposed their subjects to either stuffed toy kangaroos and koalas (associated with Australia) or kiwis (which are from New Zealand) prior to the perception task. As a result, the subjects' perception of /ɪ/ shifted depending on which toys they had seen. Therefore, Hay and Drager argued that the effect of dialect labeling is so robust that even implicit mention of another dialect would orient listeners' perception towards that dialect.

In the aforementioned socioperceptual investigations, researchers have mainly focused on the perceptual integration of linguistic and dialect information conveyed by vowels. This may be due to the fact that in languages like English, vowel differences across dialects outnumber consonant differences, and therefore are particularly important for distinguishing between regional accents. In comparison, in Mandarin Chinese, vowels remain largely constant, whereas the repertoire of consonants varies across dialects. Therefore, the current study extended the investigation of dialect labeling effects on vowel perception to consonant perception. The target feature under study is the distinction between the alveolar /s, ts, tsʰ/ and retroflex /ʂ, tʂ, tʂʰ/ sibilants in Mandarin, a place feature that is most often used for characterizing dialects of Mandarin. Two dialects of Mandarin Chinese, Beijing Mandarin and Taiwan Mandarin, were chosen to explore the biasing effects of dialect information on the perception of the /ʂ-s/ continuum. While to the best of my knowledge there is no previous study of dialect labeling effects on the perception of fricative place contrasts in other languages to make a prediction for the current study, researchers (e.g., Johnson 1991, Strand 1999, Munson 2011) have found that perceived gender affects perception of English /s/ and /ʃ/. These researchers have attributed gender effects on frication perception to exemplars or stereotypes about gender being activated—the same account that Hay, Nolan & Drager (2006) and Hay

& Drager (2010) hold for the dialect effects found on vowel perception. Therefore, dialect information was assumed to similarly affect Mandarin fricative perception.

2. Factors conditioning the effects of dialect information in the Mandarin alveolar-retroflex perception

The presence of dialect labeling effects in perceiving a consonantal place contrast is premised on dialectally governed variability in production and perception that listeners can exploit or accommodate. A recent cross-dialectal study by Chang & Shih (2015), where they compared the alveolar and retroflex sibilant produced by Beijing and Taiwan Mandarin speakers, found that while alveolar productions were not realized significantly differently between the two dialects in any vowel context, retroflex productions in the /a, i/ contexts in Beijing Mandarin had significantly lower spectral center of gravity than in Taiwan Mandarin. Upon the comparison of the alveolar-retroflex contrast, measured in terms of the spectral distance between the two sibilant categories, Beijing Mandarin exhibited a significantly larger spectral distance between the two categories than Taiwan Mandarin in the /a/ context, but not the /i/ or /u/ context.[1] Cross-dialectal variability in the alveolar-retroflex contrast has also been found in perception. Chang, Shih & Allen (2013) investigated the alveolar-retroflex perceptions in different vowel contexts by using listeners from Beijing and Taiwan. The results indicate that the Beijing and Taiwanese listeners have different perceptual boundaries along the /ʂa-sa/ and /ʂu-su/ continua, with the Beijing listeners' perceptual boundary located one step closer to the retroflex end of the continuum than the Taiwanese listeners. That is, given the same retroflex-alveolar continuum, fewer tokens were perceived as retroflexes by the Beijing listeners. The vowel context had an effect on the location of the category boundary for both Beijing and Taiwanese listeners: the perceptual boundary was located at lower frequencies along the /ʂu-su/ continuum than the /ʂa-sa/ continuum. Taken together, the cross-dialect production and perception variability of the Mandarin alveolar-retroflex contrast warrants the current study to investigate whether the interaction of linguistic and dialect

[1] Literature on Mandarin phonology (e.g., Duanmu 2000, Lin 2007) generally describes a well-distinguished and consistent alveolar-retroflex contrast for Beijing Mandarin, and various degrees of the alveolar-retroflex distinction for Taiwan Mandarin. As a result, Taiwan Mandarin has drawn much more attention than Beijing Mandarin in acoustic investigations of alveolar-retroflex production variability. Chang & Shih (2015) thus conducted a detailed comparison of the alveolar and retroflex productions in Beijing and Taiwan Mandarin as a function of vowel context and focal prominence—two factors previously studied only in Taiwan Mandarin. Nevertheless, it should be noted that Chang & Shih specifically studied the speakers who were judged to be capable of making such a place distinction; the Taiwan Mandarin subjects with merged alveolar and retroflex categories (i.e., the retroflexes are merged into their alveolar counterparts) were excluded from Chang & Shih's data analysis. That being said, Chang & Shih's findings do not generalize to the population of Taiwan Mandarin speakers who make no alveolar-retroflex distinction.

information found for vowel perception is also present in the perception of fricatives. In the case of an effect of dialect labeling on the perception Mandarin /s/ and /ʂ/, our Taiwanese listeners' categorical boundary was expected to shift towards the retroflex end of the acoustic continua regardless of vowel contexts when provided with the Beijing label (in light of the cross-dialect perception results reported in Chang, Shih & Allen 2013). On the other hand, Beijing listeners' perceptual boundary was expected to shift towards the alveolar end when provided with the Taiwan label.

In addition to the cross-dialectal variation in production and perception of a linguistic feature, listeners' stereotypes of a given dialect have also been suggested to condition the presence and/or the magnitude of dialect labeling effects on speech perception (Niedzielski 1999, Hay, Nolan & Drager 2006, Hay & Drager 2010). When a feature is commonly associated with or reinforced by social stereotypes (for example, using the realization of /ɪ/, as opposed to /æ/ and /ɛ/, to distinguish between Australian English and New Zealand English), this feature is more subject to dialect labeling manipulations. Therefore, if Beijing and Taiwanese listeners believe the major difference between their own dialect and the non-native dialect to lie in other features such as tonal realizations, they will not be subject to dialect labeling manipulation in the alveolar-retroflex perception. Accordingly, an open-ended question was included in the participants' language background questionnaire to discover the role of stereotyped knowledge of dialectal variation in fricative perception: "What is the most salient phonetic difference between Beijing Mandarin and Taiwan Mandarin to you?" Individuals that considered the alveolar-retroflex contrast to be the most salient dialect-distinguishing feature were expected to be more subject to dialect manipulation.

In light of literature that reports speech accommodation being modulated by group identity or attitudes (e.g., Bourhis, Giles & Lambert 1975, Selting 1985, Trudgill 2008, Babel 2010), listeners' attitudes towards other dialects may also play a role in the current perceptual study. Maintaining the alveolar-retroflex contrast in speech is indexed as standard pronunciation and is associated with a higher education level in both China and Taiwan. Given this feature's social indexical value, speakers have been found to exhibit stylistic retroflexing in speech: Jeng (2006) reported that retroflexion occurs more often in response to a more formal register for Taiwan Mandarin speakers; and Chung (2006) pointed out that Taiwan Mandarin speakers are more subject to hypercorrect speech (the alveolars being replaced with their retroflex counterparts) when trying to display respect or formality to an interlocutor. On the other hand, attitudes towards "standardness" may inhibit such stylistic accommodation: Liao (2010) reported that some Taiwan Mandarin speakers choose not to accommodate to, or even to distance themselves from, the more standard forms of Mandarin so as to emphasize their Taiwanese identity. Together, speech convergence (i.e., reduction of the speech differences between two dialects) or divergence (i.e., accentuation of speech differences between the two dialects) may occur to Taiwanese listeners as a result of their socio-cognitive association or dissociation with the Beijing dialect. Therefore, another open-ended question was included in the participants' language background questionnaire to elicit listeners' attitudes towards the other dialect: "How do you feel about Beijing Mandarin?" (or Taiwan Mandarin, if the background questionnaire was intended for Beijing participants.) Individuals who expressed negative feelings about the other dialect were expected to be more subject to the current dialect manipulation when given the non-native dialect label.

Since the current study specifically investigates the effects of social-indexical information conveyed by the alveolar and retroflex fricatives on perception, it is critical to control other factors that reveal dialectal information. One such factor is the auditory stimuli's tone. Fon & Chiang (1999) found that tone production in Taiwan Mandarin, compared to Beijing Mandarin, is characterized by a narrower tonal range, lower tonal heights and flatter tonal contours. As such, the realizations of the auditory stimuli's tone could provide an additional cue for identification. Whereas the contour tones (i.e., Mandarin Tones 2, 3, and 4) are characterized by the interaction of F0 direction and height, which better facilitates speaker normalization (see Wong & Diehl 2003, Lee, Lee & Shr 2011), Mandarin Tone 1 contains only one acoustic correlate (i.e., F0 height) for perception. Therefore, presenting Tone 1 tokens in isolation (such that external F0 information is not available to listeners) would minimize cues related to dialect-distinguishing tonal realizations during fricative identification.

Finally, in the case of no evidence for any influence of dialect labels, dialect stereotypes and dialect attitudes on frication perception, it is critical to ensure that listeners are indeed sensitive to the phonetic differences between the Beijing and Taiwanese dialects; that is, listeners could at least make reliable dialect categorization judgments based on the available information in the speech signal. To this end, a dialect categorization task, where listeners heard a stimulus and selected the speaker's dialect background (Beijing or Taiwan), was also included in this study.

3. Methodology

3.1 Participants

The participants were 60 Beijing Mandarin speakers (14 males, 46 females) and 60 Taiwan Mandarin speakers (17 males, 43 females). All the Beijing participants were born and raised in Beijing or its vicinity; all the Taiwanese participants were born and raised in Taiwan.[2] The age of both groups of participants ranged from 18 to 30 years old. Almost all the Beijing participants spoke only one dialect of Chinese, namely Beijing Mandarin, except for six who also spoke one other Chinese dialect. Two-thirds of the Taiwanese participants spoke Taiwanese (a southern Chinese dialect) at various levels of proficiency. None of the participants self-reported any past or present speech or hearing disorders.

3.2 Stimuli

Task 1 was an identification task (see Appendix 1 for a sample response page in Task 1), and the stimuli used in Task 1 were identical to those used in Chang, Shih & Allen (2013). They included one 8-step /ʂa-sa/ (沙-撒) continuum and one 8-step /ʂu-su/ (書-蘇) continuum. The two continua were constructed based on the alveolar-retroflex production of one male Beijing Mandarin speaker.[3] First, the frication region between the alveolar and retroflex (i.e., the region between the lowest bound of the wideband noise for /ʂ/ and /s/) was divided into 7 equal intervals on a log scale, yielding 8 cutoff frequencies (see Table 1 for the cutoff frequencies of each step of the /ʂa-sa/ and /ʂu-su/ continua). As the spectral contrast between /ʂu/ and /su/ is smaller than that between /ʂa/ and /sa/, the interval on the /ʂu-su/ continuum was smaller than that on the /ʂa-sa/ continuum, reflecting different magnitudes of the alveolar-retroflex contrast across vowels. To create an 8-step acoustic continuum from, the frication noise was edited out with a high-pass filter, starting at the cutoff frequency of /ʂa/ and /ʂu/ in steps, until reaching the cutoff frequency of /sa/ and /su/.[4] Further details of the stimulus construction are described in Chang, Shih & Allen (2013). Two 8-step continua /i-y/ (衣-淤) and /tʰi-ti/ (梯-滴) were also included as fillers. All the stimuli carried Mandarin Tone 1.

---

[2] Extensive information about the participants' city of birth was not collected, as the current study only required that the participants were respectively from the greater Beijing area or Taiwan.

[3] Since Chang & Shih (2015) found that the alveolar productions of the two dialects did not significantly differ in terms of spectral center of gravity in any vowel context, while the retroflex productions in the /a, i/ contexts in Beijing Mandarin had significantly lower spectral center of gravity, Beijing Mandarin speakers' acoustic contrast between /ʂ/ and /s/ is more likely to be greater than that in Taiwan Mandarin speech. To maximize the acoustic contrast for subsequent /ʂ-s/ continuum construction, Chang, Shih & Allen (2013) used the speech produced by one Beijing Mandarin speaker. It may be questioned whether listeners could identify the speaker as a Beijing Mandarin speaker. This possibility was tested with a dialect classification experiment (see Task 2). The results of Task 2 showed that Beijing and Taiwanese listeners' dialect categorization performance based on the alveolar- and retroflex-initial monosyllabic words which bear Tone 1 was no better than chance level. Therefore, the choice of stimuli for Task 1 should not confound the identification results.

[4] The two continua are available for public download at https://drive.google.com/file/d/0B0QGao8jtzMLZmg3c0RVY1pVN2s/edit?usp=sharing.

Table 1. Cutoff frequencies (kHz) of each step of the /ʂa-sa/ and /ʂu-su/ continua

continuum steps	/ʂa-sa/ continuum	/ʂu-su/ continuum
Step 1	1.39	1.24
Step 2	1.61	1.4
Step 3	1.89	1.61
Step 4	2.24	1.85
Step 5	2.67	2.12
Step 6	3.15	2.41
Step 7	3.68	2.72
Step 8	4.23	3.12

Note: The cutoff frequency of each step was located using the logarithmic interpolation formula[5] in Deserno (2004).

As addressed in Section 2, if the Beijing and Taiwanese listeners could not reliably distinguish between the two regional accents to begin with, they would not be subject to the manipulation of dialect information in fricative perception, thereby rendering null results. One way to rule out this possibility is to test whether listeners are able to categorize speakers by dialect. Since the stimuli used in Task 1 (i.e., the /ʂa-sa/ and /ʂu-su/ continua) carried Tone 1, including stimuli that carry other Mandarin tones may allow discussion of contribution of tonal realizations to dialect classification. As a result, the monosyllabic stimuli[6] used in Task 2 respectively carried each of the four Mandarin lexical tones (see Appendix 2 for a sample response page in Task 2 and Appendix 3 for the list of the monosyllabic stimuli). An additional set of stimuli—the first two sentences of the Chinese version of *The North Wind and the Sun*—were also included in Task 2 to compare categorization performances based on single-word and sentential stimuli (see Appendix 4 for the sentence stimuli). The monosyllabic and sentence stimuli used in Task 2 were produced by eight Beijing Mandarin speakers and eight Taiwan Mandarin speakers.

[5] In the formula $y=y_2^f y_1^{1-f}$, the cutoff frequency of a given stimulus step corresponds to y, which is located between two ticks y_1 and y_2 on the y-axis of the AI-grams (see Chang, Shih & Allen 2013), and f is the distance between y and y_1 over the distance between y_1 and y_2.

[6] The alveolar-retroflex pairs in Tasks 1 and 2 were adopted from Chang and Shih (2015) and were all controlled for word frequency (between the log frequencies of 3.4 and 5). See Chang and Shih (2015) for details on how the lexical frequencies of these syllables were derived. Additionally, it should be noted that the alveolar-retroflex pairs used in Task 2 were aspirated and unaspirated affricates, instead of fricatives, for the consideration of word-frequency control.

3.3 Procedure

Prior to the experiment, the participants were randomly divided into two groups, where they were given different information about the speaker's dialect background (i.e., Beijing or Taiwan). The participants were seated at a computer in a quiet lab space at Beijing Language and Culture University and National Taipei University of Technology respectively. They signed an informed consent form approved by the Research Ethic Office of National Taiwan University.

Task 1 was a two-alternative forced-choice identification task. Before each sound was played, a page of text appeared on the computer screen, instructing the participants that the speaker was from Beijing (or Taiwan, depending on the grouping assignment). Upon hearing one stimulus, listeners had to identify the sound by choosing one of the two characters representing either syllable from the minimal pair (namely, /ʂa/ 沙 vs. /sa/ 撒; /ʂu/ 書 vs. /su/ 蘇; /tʰi/ 梯 vs. /ti/ 滴; /i/ 衣 vs. /y/ 淤) (see Appendix 1 for a sample response page). Before the real task, there was a three-trial practice with non-alveolar and non-retroflex stimuli produced by a different speaker. The participants were told that the task was timed and were encouraged to respond as quickly as possible. Task 2 was a two-alternative forced-choice dialect categorization task. In the first part of Task 2, listeners were instructed to judge whether the speaker was from Beijing or Taiwan based on his or her pronunciation of the character (i.e., a monosyllabic word) shown on the computer screen (see Appendix 2 for a sample response page). Each sound was played two seconds after the corresponding text page appeared. The second part of Task 2 followed the same procedure, except that the stimuli were sentences. A three-trial practice, with stimuli produced by different speakers from those at the testing stage, was given before proceeding to each part of Task 2. The participants were asked to respond to each trial as quickly as possible. The presentation and randomization of the stimuli were performed by using E-Prime (v.2.0) for both Tasks 1 and 2, and all responses were automatically logged.

At the end, the participants completed a questionnaire which covered their language backgrounds and the previously mentioned two open-ended questions about dialect attitudes and dialect stereotypes. All participants received cash compensation for their participation. Given the different orthographic systems used in China and Taiwan, all printed materials and experimental instructions on the computer monitor came in two versions, i.e., simplified and traditional Chinese, and were distributed to the participants based on their place of origin. As the dialect of the experimenter can result in participants' accommodation in production (Hay, Drager & Warren 2009) as well as in perception (Hay, Nolan & Drager 2006), the experiment in Beijing was conducted by an experimenter who spoke Beijing Mandarin, while the experiment in Taiwan was conducted by two Taiwan Mandarin speakers.

4. Results

4.1 Effects of dialect information on the Mandarin alveolar-retroflex perception

Figures 1 and 2 show the identification results from Task 1. Mean retroflex identification responses were plotted against steps along the /ʂa-sa/ and /ʂu-su/ continua across dialect labeling conditions (i.e., the "Beijing" label or the "Taiwan" label) for each group of listeners. Using the lme4 package (Bates & Maechler 2010) in the R software (R Development Core Team 2010), a mixed-effects logistic regression analysis was performed. The identification data were binary coded, and therefore logistic regression analysis was an appropriate statistical method. The analysis was performed on the /ʂa-sa/ and /ʂu-su/ identification data separately, since the steps along the /ʂa-sa/ continuum correspond to different frequency cutoffs from those along the /ʂu-su/ continuum (see Table 1). The variables of stimulus step and dialect labeling were added as fixed effects, and participants were added as a random intercept effect.

The analysis of the Beijing listeners' data started with a by-subject intercept-only model and then added the fixed factors (stimulus step, dialect labeling) one by one. The interaction term was subsequently added to the model. The factors and factor interactions were retained if the model was significantly improved. Regarding the /ʂa-sa/ identification (see Figure (1a)), the results of the best-fitting model (see Table 2) indicated a significant effect of the stimulus step ($Wald\ \chi^2(1) = 369.9$, $p < .001$). Regarding the /ʂu-su/ identification (see Figure (1b)), the best-fitting model (see Table 3) also included only the stimulus step ($Wald\ \chi^2(1) = 345.8$, $p < .001$).

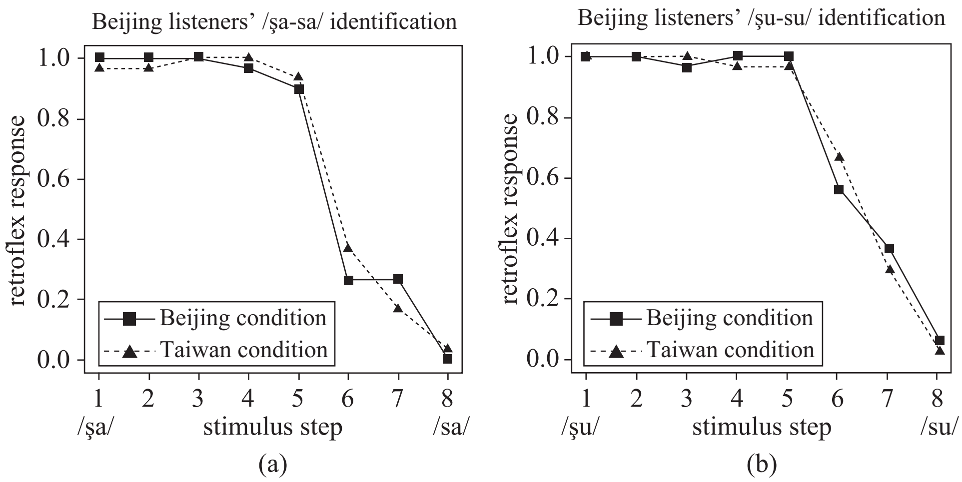

Figure 1. The Beijing listeners' retroflex responses to tokens from the /ʂa-sa/ and /ʂu-su/ continua across dialect labeling conditions

Table 2. Summary of the best-fitting mixed-effects logistic model for the Beijing listeners' /ʂa-sa/ identification (* p < .001)**

Effects	Measures		
	Estimated Coefficient	Std. Error	z-value
(Intercept)	10.745	1.297	8.283***
Stimulus Step	-1.814	0.217	-8.351***

Table 3. Summary of the best-fitting mixed-effects logistic model for the Beijing listeners' /ʂu-su/ identification (* p < .001)**

Effects	Measures		
	Estimated Coefficient	Std. Error	z-value
(Intercept)	16.993	2.648	6.418***
Stimulus Step	-2.599	0.404	-6.439***

Regarding the Taiwanese listeners' /ʂa-sa/ identification (see Figure (2a)), the best-fitting model (see Table 4) included only stimulus step (*Wald* $\chi^2(1) = 205.5$, $p < .001$). For the /ʂu-su/ identification (see Figure (2b)), the results of the best-fitting model (see Table 5) only indicated a significant main effect of the stimulus step (*Wald* $\chi^2(1) = 167.55$, $p < .001$).

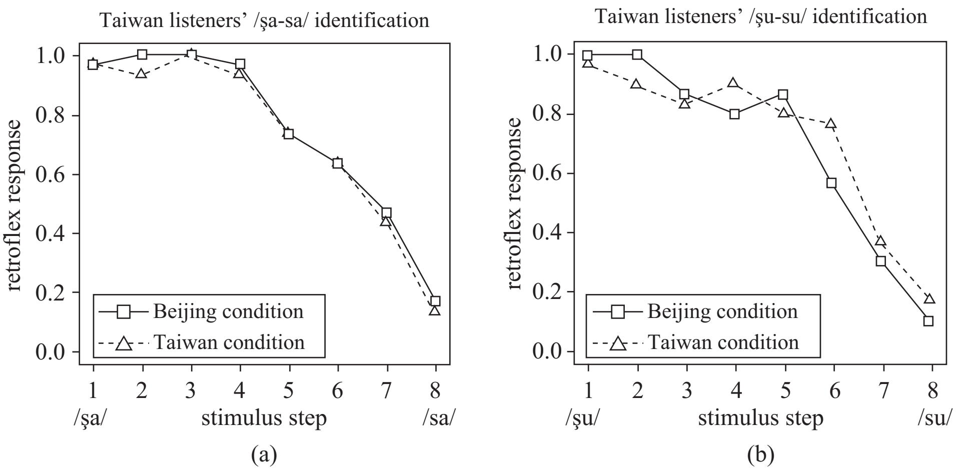

Figure 2. The Taiwanese listeners' retroflex responses to tokens from the /ʂa-sa/ and /ʂu-su/ continua across dialect labeling conditions

Table 4. Summary of the best-fitting mixed-effects logistic model for the Taiwanese listeners' /ʂa-sa/ identification (* p < .001)**

Effects	Measures		
	Estimated Coefficient	Std. Error	z-value
(Intercept)	6.684	0.698	9.582***
Stimulus Step	-1.006	0.109	-9.258***

Table 5. Summary of the best-fitting mixed-effects logistic model for the Taiwanese listeners' /ʂu-su/ identification (*p < .001)**

Effects	Measures		
	Estimated Coefficient	Std. Error	z-value
(Intercept)	6.253	0.655	9.555***
Stimulus Step	-0.894	0.100	-8.932***

To sum up, the information on the speaker's dialect background was found to have no effects on either groups of listeners' decisions along the /ʂa-sa/ and /ʂu-su/ continua. That the identification functions were affected by stimulus steps indicates that fewer retroflex percepts were elicited later in the continua. The dialect-conditioned variation in perception found here was in line with Chang, Shih & Allen (2013) findings: the location of the /ʂ-s/ perceptual boundary based on the 50% crossover point differed between Beijing and Taiwanese listeners. The Beijing listeners' perceptual boundary along the /ʂa-sa/ and /ʂu-su/ continua was located one step closer to the retroflex end of the continuum than the Taiwanese listeners', regardless of dialect labeling conditions.

4.2 Roles of dialect stereotypes and dialect attitudes in the Mandarin alveolar-retroflex perception

As stated in Section 2, two open-ended questions were included in the participants' language background questionnaire. The first question explicitly asked what phonetic feature was considered by the participants to best distinguish between Beijing Mandarin and Taiwan Mandarin. The second question was about the participants' attitudes towards the Mandarin dialect that was not native to them.

The questionnaire data were coded according to the categories that were observed to dominate the data. For the first question, the Beijing participants' responses were categorized into 1) retroflex articulation (including the alveolar-retroflex contrast and *er*-suffixation), 2) tone and intonation, 3) clarity of speech, 4) voice quality, and 5) others (i.e., answers that only occurred once and could not be categorized under the previous four categories). The categories which comprise the Taiwan participants' responses were almost the same as their Beijing counterparts, except that voice quality was never mentioned by the Taiwan participants and speech rate was a recurring theme only in Taiwan participants' responses. See Table 6 for the distribution of the coded results for the first question.

Table 6. Distribution (raw counts/percentage) of most salient dialect-distinguishing phonetic features considered by the Beijing and Taiwanese listeners

Beijing (n = 60)		Taiwan (n = 60)	
Retroflex articulation %	27/45	Retroflex articulation %	36/60
Tone and intonation %	15/25	Tone and intonation %	11/18.3
Others %	9/15	Others %	7/11.7
Clarity of speech %	5/8.3	Clarity of speech %	4/6.7
Voice quality %	4/6.7	Speech rate %	2/3.3

As for the second question, both Beijing and Taiwanese participants' responses were categorized into positive, neutral and negative. In responses categorized as neutral, both participants stated how the non-native dialect was similar to (or different from) their own dialect, or they simply claimed that they did not have any specific feeling about the non-native dialect. In responses considered positive, the Beijing participants described the Taiwanese dialect as gentle, soft or pleasant, whereas the Taiwanese participants described the Beijing dialect to be standard, correct or articulate. In responses deemed negative, the Taiwanese dialect was considered unpleasant or difficult to be recognized, while the Beijing dialect was described as unfriendly or too retroflexed. The distribution of the coded responses to the second question is displayed in Table 7.

Table 7. Distribution (raw counts/percentage) of the Beijing listeners' and Taiwanese listeners' attitudes towards each other's dialect

Beijing (n = 60)		Taiwan (n = 60)	
Positive %	24/40	Positive %	16/26.7
Neutral %	18/30	Neutral %	29/48.3
Negative %	18/30	Negative %	15/25

To test whether stereotypes of and attitudes towards regional accent differences between Beijing Mandarin and Taiwan Mandarin affected listeners' identification decisions, these two variables were added as fixed factors to the linear mixed-effects models introduced in Section 4.1. The interaction terms were subsequently added to the models. The results showed that the addition of the two factors and their interaction terms did not improve the model of either Beijing or Taiwanese listeners' data, suggesting that listeners' stereotyped knowledge and attitudes of the dialectal variation did not influence their fricative identification in Task 1.

4.3 Dialect categorization based on monosyllabic words and sentences

Task 2 examined whether the Beijing and Taiwanese listeners could distinguish between the two regional accents. In classifying dialects based on short sentences, both groups achieved a high level of accuracy: the Beijing participants exhibited 92.8% (SD = 25.9) accuracy in correctly categorizing the 16 speakers by dialect and the Taiwanese participants exhibited 92.5% (SD = 26.4) accuracy. Based on monosyllabic words, the Beijing participants correctly categorized the 16 speakers by dialect at 65.8% (SD = 48.1) and the Taiwanese participants did so at 62.7% (SD = 48.4). Upon closer inspection, dialect categorization of monosyllabic words varied considerably in accuracy when being further divided by the tokens' place of articulation (i.e., alveolar- or retroflex-initial) and tone (i.e., the four Mandarin tones). Table 8 displays the descriptive statistics for the Beijing and Taiwanese listeners' dialect categorization accuracy of monosyllabic alveolar- and retroflex-initial tokens across the four lexical tones.

Table 8. The Beijing and Taiwanese listeners' mean accuracy (in % with SD) of dialect classification for monosyllabic alveolar- and retroflex-initial tokens across tones

	Beijing listeners		**Taiwanese listeners**	
	Alveolar t(59); p value	**Retroflex** t(59); p value	**Alveolar** t(59); p value	**Retroflex** t(59); p value
Tone 1	53.1 (13.8) 1.71 (p = .092)	52.4 (13.1) 1.39 (p = .17)	50.4 (12.2) 0.26 (p = .79)	52.8 (12.8) 1.68 (p = .09)
Tone 2	62.8 (12.1) 8.16 (p < .001)	76.1 (10.5) 19.33 (p < .001)	60.8 (14.4) 5.82 (p < .001)	69 (15.1) 9.75 (p < .001)
Tone 3	62.9 (13.2) 7.56 (p < .001)	76 (10.1) 19.98 (p < .001)	58.5 (15.8) 4.14 (p < .001)	70 (16.2) 9.58 (p < .001)
Tone 4	63.1 (14.1) 7.18 (p < .001)	80 (9.8) 23.6 (p < .001)	63.3 (15.6) 6.63 (p < .001)	76.4 (12.6) 16.26 (p < .001)

Note: Results from t-tests evaluating whether the identification accuracy was above the chance level (50%) were also included.

To confirm whether dialect categorization accuracy was affected by place of articulation and tone of the monosyllabic tokens, the accuracy data were analyzed with a mixed-effects logistic regression model, with place and tone as the fixed effects and participants as a random intercept effect. The analysis started with an intercept-only model and then added the fixed factors one by one. The factors and factor interactions were retained if the model was significantly improved. Regarding the Beijing listeners' data, the best-fitting model (see Table 9) included place (*Wald* $\chi^2(1) = 90.097$, $p < .001$), tone (*Wald* $\chi^2(3) = 88.455$, $p < .001$) and the place × tone (*Wald* $\chi^2(3) = 13.172$, $p < .01$) interaction. Tukey-adjusted pairwise comparisons of estimated regression coefficients and z-statistics showed the following results for Beijing listeners:

1) Tone 1 alveolar-initial tokens had significantly lower dialect categorization accuracy than the other three tones (all at $p < .01$) whereas accuracy based on Tones 2, 3 and 4 alveolar-initial tokens did not significantly differ from one another.

2) Tone 1 retroflex-initial tokens had significantly lower dialect classification accuracy than the other three tones (all at $p < .001$) whereas accuracy based on Tones 2, 3 and 4 retroflex-initial tokens did not significantly differ from one another.

3) Dialect categorization accuracy based on Tone 1 retroflex-initial tokens was not significantly different from that based on Tone 1 alveolar-initial tokens. On the other hand, Tones 2, 3 and 4 retroflex-initial tokens had a significantly higher level of accuracy than their respective alveolar-initial counterparts (all at $p < .001$).

Table 9. Summary of the best-fitting mixed-effects logistic model for the Beijing listeners' dialect categorization performance based on monosyllabic words (*p < .05; ***p < .001)

Effects	Measures		
	Estimated Coefficient	Std. Error	z-value
(Intercept)	0.123	0.077	0.112
Place Retroflex	0.273	0.107	2.556*
Tone 2	0.402	0.108	3.740***
Tone 3	0.409	0.108	3.794***
Tone 4	0.414	0.108	3.848***
Place Retroflex: Tone 2	0.298	0.157	1.894
Place Retroflex: Tone 3	0.263	0.157	1.672
Place Retroflex: Tone 4	0.583	0.162	3.605***

Regarding the Taiwanese listeners' data, the best-fitting model (see Table 10) included place (*Wald* $\chi^2(1) = 48.93$, $p < .001$), tone (*Wald* $\chi^2(3) = 107.57$, $p < .001$), as well as the place × tone (*Wald* $\chi^2(3) = 11.87$, $p < .01$) interaction. Tukey-adjusted pairwise comparisons indicated the following three results for the Taiwanese listeners:

1) Tone 1 alveolar-initial tokens had a significantly lower level of dialect categorization accuracy than the other three tones (all at $p < .01$), whereas accuracy based on Tones 2, 3 and 4 alveolar-initial tokens did not significantly differ from one another.

2) Tone 1 retroflex-initial tokens had a significantly lower level of dialect classification accuracy than Tones 2, 3 (both at $p < .01$) and 4 ($p < .001$). Although there was no significant accuracy difference between Tones 3 and 4, or Tones 2 and 4, Tone 2 rendered a significantly lower level of accuracy than Tone 4 ($p < .05$).

3) Dialect categorization accuracy based on Tone 1 retroflex-initial tokens was not significantly different from Tone 1 alveolar-initial tokens. On the other hand, Tones 2, 3 and 4 retroflex-initial tokens had a significantly higher level of classification accuracy than their respective alveolar-initial counterparts (all at $p < .001$).

Table 10. Summary of the best-fitting mixed-effects logistic model for the Taiwanese listeners' dialect categorization performance based on monosyllabic words (*p < .05; **p < .01; ***p < .001)

Effects	**Measures**		
	Estimated Coefficient	Std. Error	z-value
(Intercept)	0.017	0.085	0.202
Place Retroflex	0.131	0.107	1.227
Tone 2	0.433	0.108	4.015***
Tone 3	0.333	0.107	3.100**
Tone 4	0.542	0.109	4.988***
Place Retroflex: Tone 2	0.205	0.155	1.324
Place Retroflex: Tone 3	0.385	0.155	2.484*
Place Retroflex: Tone 4	0.509	0.159	3.195**

To evaluate whether the accuracy of identification was above the chance level of 50%, sixteen t-tests (2 dialects × 2 places of articulation × 4 tones) were conducted on the identification accuracy data drawn from monosyllabic stimuli. A Bonferroni correction was used for controlling for Type 1 errors as a result of multiple comparisons. The results, as shown in Table 8, indicated that both Beijing and

Taiwanese listeners were able to identify speakers' dialect background above the chance level for monosyllabic words carrying a contour tone. In contrast, both groups identified the level-tone (Tone 1) tokens at the chance level.

To sum up, both Beijing and Taiwanese participants were able to classify speakers by dialect with different degrees of accuracy. Noticeably, a higher level of dialect categorization accuracy was achieved with the sentence material than with monosyllabic tokens. Dialect categorization accuracy based on monosyllabic words was the highest for retroflex-initial tokens carrying contour tones, followed by alveolar-initial tokens carrying contour tones; the accuracy was lowest and at the chance level for alveolar- and retroflex-initial tokens carrying the level tone.

5. Discussion

In view of the previous studies on the influence of dialect labeling on vowel perception, this study investigated whether the sensitivity of the perceptual system to dialect information is also present in fricative perception. By means of the Mandarin alveolar-retroflex contrast, which is a feature often used for characterizing cross-dialect variation between Beijing Mandarin and Taiwan Mandarin, the current study specifically examined whether the location of the perceptual boundary over the retroflex-alveolar continua would shift when given a different dialect label. The results of Task 1 revealed similar perceptual variation as a result of dialect and vowel context as reported in Chang, Shih & Allen (2013). However, the alveolar-retroflex identification judgments were not affected by the dialect labels within either groups of listeners.

A further analysis was carried out on whether listeners' dialect stereotypes and attitudes could explain the null dialect labeling effects as these are the two factors which may condition speech accommodation. The analysis assumes that people who held a stereotype regarding retroflexion in speech and who held negative attitudes towards the other dialect were expected to be more subject to dialect labeling manipulation. However, the statistical results did not meet these expectations. That is, dialect stereotypes and attitudes were found to have no influence on listeners' identification decisions.

The null results of Task 1 were surprising, given that the factors under study (i.e., dialect labeling, dialect stereotypes, and dialect attitudes) have all been reported to have biasing effects on vowel perception (see studies reviewed in Sections 1 and 2). A factor which may have contributed to the null results was the lack of control of the listeners' capabilities to distinguish between alveolar and retroflex sibilants in their own production.[7] Fridland & Kendall (2012) and Kendall & Fridland (2012) closely examined how an individual's production of a vowel contrast was related to his/her perception of the same contrast. They found that *cross-regional* as well as *intra-regional* variation in production was accompanied by differences in perception. Therefore, controlling the listeners in terms of whether or not and to what extent they make the alveolar-retroflex contrast in production is critical to future investigations of dialect labeling effects in the alveolar-retroflex perception.

One may also suspect that the null results have to do with the listeners' lack of knowledge of dialectal variation regarding the alveolar-retroflex contrast. However, the results of Task 2 showed that given sentence stimuli, both groups of listeners were able to identify 16 speakers by dialect with over 90% accuracy. In addition, listeners were able to identify the speakers' dialect background above the chance level for monosyllabic alveolar- and retroflex-initial tokens that carry contour tones, although dialect categorization performance based on the alveolar- and retroflex-initial tokens which bear the level Tone 1 was at the chance level. Therefore, the null results in Task 1 should not be due to listeners' not being able to distinguish between the two regional accents.

Alternatively, the null results in Task 1 may relate to the choice of the stimuli's tone. As stated in Section 2, contour tones, which have the information of F0 direction in addition to F0 height, may better facilitate the process of speaker normalization than the level tone, thereby assisting with dialect categorization. Indeed, Task 2 found that both groups of listeners' dialect classification performance based on monosyllabic tokens which bear the level Tone 1 was at the chance level while their contour-tone token performance was significantly above the chance level. Now, since the stimuli used in Task 1 all carried Tone 1 (which provided fewer dialect-distinguishing cues than the realizations of contour tones), the absence of dialect labeling effects may be attributed to the insufficient cues contained in fricatives (i.e., only place-distinguishing cues available) to activate exemplars indexed with the respective dialects under study.

In light of the findings of the current study so far, one question that follows is whether fricatives contain less social-indexical information than vowels. But this should not be the case given the number of studies that show perceived gender affecting the perception of the English /s-ʃ/ continuum (e.g., Johnson 1991, Strand 1999, Munson 2011). To the best of my knowledge, no previous studies have shown

[7] The realizations of the alveolar and retroflex productions, particular in Taiwan Mandarin, have been found to be subject to factors of gender and region of birth (e.g., Chuang & Fon 2010). Therefore, inclusion of the alveolar-retroflex production data would have enriched the discussion of whether these social variables contribute to the effects of dialect labeling in perception.

the effects of perceived dialect on the perception of English fricatives or the effects of perceived gender on the perception of Mandarin fricatives. Therefore, it is impossible to speculate at this point whether dialect information is stored or represented qualitatively differently from gender in fricative exemplars, or whether the robustness of the phonological contrast (i.e., the English /s-ʃ/ contrast being stronger than Mandarin /s-ʂ/ contrast) may account for the absence of perceptual accommodation found here.

Since Task 1 used fricative-initial tokens whereas Task 2 used aspirated and unaspirated affricate-initial monosyllabic words, the two reviewers raised the question as to whether manner of articulation influenced the null results. In other words, could dialect information have been found to bias perception if affricate-initial tokens had instead been used in Task 1? Chang and Shih (2015) similarly questioned whether acoustic measures (frication duration, syllable duration, frication amplitude, and spectral center of gravity) of alveolar and retroflex sibilants vary as a function of manners of articulation. By means of a discriminant analysis, Chang and Shih found the four acoustic measures to poorly predict accurate manner of articulation identifications, thereby suggesting that manner of articulation did not affect their acoustic measures. Therefore, affricates are unlikely to provide different acoustic cues (as far as the four acoustic measures in Chang & Shih (2015)) than fricatives and facilitate the use of dialect information in the alveolar-retroflex perception in Task 1.

Finally, despite the previous evidence regarding the biasing effects of social information on perception, studies have also found limited evidence for such effects. Lawrence (2015) replicated Niedzielski's (1999) vowel matching paradigm by using the vowels in BATH and STRUT while manipulating the dialect labels of "Sheffield, Northern England" and "London, Southern England". Despite the fact that the salience of the two vowels' different realizations reinforces the stereotypes of Northern and Southern varieties, Lawrence did not find any significant dialect labeling effects in his experiment. To explore whether social status cues would affect the perception of morphosyntactic variation in the same way as linguistic cues affect social perception, Squires (2013) had her American subjects rate the acceptability of the two subject+verb agreement constructions (NP+don't and there's+NP) in high- and low- status guises. She found that there is no evidence of social status information affecting linguistic judgments. Therefore, this study concludes that the role of social information in perception may not be as straightforward as the previous literature has suggested. What needs to be further investigated is whether the biasing effects of social information on speech perception are sensitive to types of phonological units (vowel, consonant, or tone), dependent on types of social information (e.g., gender or dialect), or conditioned by individual-specific production patterns.

Appendix 1. A response page prompting for decision in Task 1

說話者是一個台灣人，
你覺得他唸的是哪一個字？

沙	撒

Appendix 2. A response page prompting for decision (based on monosyllabic words) in Task 2

說話者說的是『擦』
你覺得他來自於哪裡？

北京	台灣

Appendix 3. List of the monosyllabic stimuli for Task 2

Alveolar-initial words (pinyin)	Retroflex-initial words (pinyin)	Fillers (pinyin)
擦 (cā)	差 (chā)	西 (xī)
足 (zú)	逐 (zhú)	伯 (bó)
子 (zǐ)	只 (zhǐ)	喊 (hǎn)
字 (zì)	至 (zhì)	過 (guò)

Appendix 4. The first two sentences of *The North Wind and the Sun*

有一次，北風跟太陽正在爭論誰的本領大。他們正好看到有個人走過，那個人身上穿了一件厚袍子。

References

Babel, Molly. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society* 39.4:437-456.

Bates, Douglas, and Martin Maechler. 2010. *lme4: Linear Mixed-Effects Models using 'Eigen' and S4*. Retrieved May 10, 2015, from http://cran.r-project.org/web/packages/lme4/index.html.

Bourhis, Richard Y., Howard Giles, and Wallace E. Lambert. 1975. Social consequences of accommodating one's style of speech: A cross-national investigation. *International Journal of the Sociology of Language* 6:55-72.

Chang, Yung-Hsiang Shawn, and Chilin Shih. 2015. Place contrast enhancement: The case of the alveolar and retroflex sibilant production in two dialects of Mandarin. *Journal of Phonetics* 50:52-66.

Chang, Yung-Hsiang Shawn, Chilin Shih, and Jont B. Allen. 2013. Variability in cross-dialect perception of the Mandarin alveolar-retroflex contrast. *Proceedings of the International Conference on Phonetics of the Languages in China*, ed. by Wai-Sum Lee, 115-118. Hong Kong: City University of Hong Kong.

Chuang, Yu-Ying, and Janice Fon. 2010. The effect of prosodic prominence on the realizations of voiceless dental and retroflex sibilants in Taiwan Mandarin spontaneous speech. *Proceedings of Speech Prosody 2010*. Chicago: Doubletree Hotel Magnificent Mile. Retrieved May 10, 2015, from http://www.speechprosody2010.illinois.edu/papers/100414.pdf.

Chung, Karen Steffen. 2006. Hypercorrection in Taiwan Mandarin. *Journal of Asian Pacific Communication* 16.2:197-214.

Clopper, Cynthia G., and David B. Pisoni. 2004. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32.1:111-140.

Clopper, Cynthia G., and David B. Pisoni. 2007. Free classification of regional dialects of American English. *Journal of Phonetics* 35.3:421-438.

Deserno, Markus. 2004. *Linear and Logarithmic Interpolation*. Retrieved March 1, 2012, from https://www.cmu.edu/biolphys/deserno/pdf/log_interpol.pdf.

Duanmu, San. 2000. *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

Evans, Bronwen. G., and Paul Iverson. 2004. Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. *The Journal of the Acoustical Society of America* 115.1:352-361.

Fon, Janice, and Wen-yu Chiang. 1999. What does Chao have to say about tones?—A case study of Taiwan Mandarin. *Journal of Chinese Linguistics* 27.1:13-37.

Fridland, Valerie, and Tyler Kendall. 2012. Exploring the relationship between production and perception in the mid front vowels of U.S. English. *Lingua* 122.7:779-793.

Hay, Jennifer, and Katie Drager. 2010. Stuffed toys and speech perception. *Linguistics* 48.4:865-892.

Hay, Jennifer, Aaron Nolan, and Katie Drager. 2006. From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review* 23.3:351-379.

Hay, Jennifer, Katie Drager, and Paul Warren. 2009. Careful who you talk to: An effect of experimenter identity on the production of the NEAR/SQUARE merger in New Zealand English. *Australian Journal of Linguistics* 29.2:269-285.

Jeng, Jing-Yi. 2006. The acoustic spectral characteristics of retroflexed fricatives and affricates in Taiwan Mandarin. *Journal of Humanistic Studies* 40.1:27-48.

Johnson, Keith. 1991. Differential effects of speaker and vowel variability on fricative perception. *Language and Speech* 34.3:265-279.

Kendall, Tyler, and Valerie Fridland. 2012. Variation in perception and production of mid front vowels in the U.S. Southern Vowel Shift. *Journal of Phonetics* 40.2:289-302.

Lawrence, Daniel. 2015. Limited evidence for social priming in the perception of the BATH and STRUT vowels. *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: University of Glasgow. Retrieved May 10, 2015, from https://www.internationalphoneticassociation.org/icphs-proceedings/ ICPhS2015/Papers/ICPHS0244.pdf.

Lee, Chao-Yang, Yuh-Feng Lee, and Chia-Lin Shr. 2011. Perception of musical and lexical tones by Taiwanese-speaking musicians. *The Journal of the Acoustical Society of America* 130.1:526-535.

Liao, Sze-Wei. 2010. Identity, Ideology, and Language Variation: A Sociolinguistic Study of Mandarin in Central Taiwan. Doctoral dissertation, University of California Davis, Davis, CA.

Lin, Yen-Hwei. 2007. *The Sounds of Chinese*. Cambridge: Cambridge University Press.

Maye, Jessica, Richard N. Aslin, and Michael K. Tanenhaus. 2008. The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science* 32.3:543-562.

Munson, Benjamin. 2011. The influence of actual and imputed talker gender on fricative perception, revisited (L). *The Journal of the Acoustical Society of America* 130.5:2631-2634.

Niedzielski, Nancy. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18.1:62-85.

Purnell, Thomas, William Idsardi, and John Baugh. 1999. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology* 18.1:10-30.

R Development Core Team. 2010. *The R Project for Statistical Computing*. Retrieved May 10, 2015, from http://www.R-project.org/.

Selting, Margret. 1985. Levels of style-shifting: Exemplified in the interaction strategies of a moderator in a listener participation programme. *Journal of Pragmatics* 9.2:179-197.

Squires, Lauren. 2013. It don't go both ways: Limited bidirectionality in sociolinguistic perception. *Journal of Sociolinguistics* 17.2:200-237.

Strand, Elizabeth A. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology* 18.1:86-100.

Trudgill, Peter. 2008. Colonial dialect contact in the history of European languages: On the irrelevance of identity in new-dialect formation. *Language in Society* 37.2:241-280.

van Bezooijen, Renée, and Charlotte Gooskens. 1999. Identification of language varieties: The contribution of different linguistic levels. *Journal of Language and Social Psychology* 18.1:31-48.

Wong, Patrick C. M., and Randy L. Diehl. 2003. Perceptual normalization for inter- and intratalker variation in Cantonese level tones. *Journal of Speech, Language, and Hearing Research* 46.2: 413-421.

[Received March 21, 2016; revised June 6, 2016; accepted October 14, 2016]

Department of English
National Taipei University of Technology
Taipei, TAIWAN
Yung-Hsiang Shawn Chang: shawnchang@ntut.edu.tw

方言資訊對中文平捲音對比之影響

張詠翔
國立臺北科技大學

社會感知文獻指出，聽者對於說話者方言背景的主觀認定，會造成聽者對相同語音有不同的感知。有鑒於這些研究僅只探討母音，本研究用台灣華語和北京普通話裡平捲音對比，來探討摩擦音感知是否也存在類似的感知調適現象。實驗中透過方言標籤（北京或台灣）的操弄，進行對聽者/ʂa-sa/和/ʂu-su/的語音連續體的辨識測驗。結果顯示，方言標籤的操弄並未影響聽者的語音辨識決定。方言既定印象和方言態度也未對辨識結果造成顯著影響。方言標籤未對中文摩擦音感知造成影響，可能跟本研究未控制聽者的平捲音產成能力有關，也有可能是中文摩擦音訊號裡欠缺辨別方言的資訊。有鑒於先前研究亦提出社會索引訊息未總是能造成語音感知調適，本研究建議未來相關研究需探討語音感知調適是否存在先決條件（如：受試者相關變因之控制、刺激材料類別之選擇）。

關鍵詞：社會索引訊息、感知偏差、中文平捲音對比、方言既定印象