The effect of dialectal variation on word recognition

A case from Taiwan Southern Min

Yu-An Lu
National Chiao Tung University

Previous studies on Chinese dialect variation have mostly focused on the description of dialects, the regions where these dialects are spoken, attitudes towards dialects, and acoustic differences across dialects. The present study draws on experimental evidence concerning a vowel difference in two Taiwan Southern Min (TSM) dialects to provide more understanding on how non-contrastive, dialectal variations may affect speakers' processing of speech. The variation of interest is a phonemic difference, [ə] and [ɔ], in the vowel inventory in two TSM dialects, in which the difference signals a lexical contrast in one dialect (e.g. [ə-a] 'oyster' vs. [ɔ-a] 'taro') but not in the other ([ɔ-a] 'oyster, taro'). A long-term repetition-priming experiment investigating the word recognition involving the two vowels revealed a dialect effect on TSM speakers' word recognition in accordance with prior exposure, native-ness and variant frequency. Implications of the findings are provided.

Keywords: word recognition, Taiwanese Southern Min, dialectal variation, repetition priming

1. Introduction

Generative phonologists who assume the principle of economy posit that words are represented in the lexicon in the format of abstract phonological representations that encode only contrastive information from the surface variations (e.g. Chomsky & Halle 1968; Kenstowicz & Kisseberth 1979; Clements 2003). In this classic view of phonological representations, variations are generally treated as peripheral. Recent research, however, has shifted its attention to investigate variations in speech, such as allophonic variations (e.g. Boomershine et al. 2008; Babel & Johnson 2010; Mackenzie et al. 2014; Garcia & Campos-Astorkiza 2015), free

https://doi.org/10.1075/lali.00048.lu
Language and Linguistics 20:4 (2019), 535–568
ISSN 1606-822X / E-ISSN 2309-5067

variations (e.g. Sumner & Samuel 2005; Ranbom & Connine 2007), dialectal variations (e.g. Sumner & Samuel 2009), social indexical variations (e.g. Babel & Russell 2015; Chuang 2017), and frequency-induced variations (e.g. Patterson & Connine 2001) and has found experimental support that non-contrastive elements may also have an impact on how speakers process speech.

The present study draws on experimental evidence concerning a vowel difference in two Taiwan Southern Min (TSM) dialects to provide better understanding on how non-contrastive, dialectal variations may affect speakers' processing of speech. Specifically, this study seeks to understand how dialectal differences affect word recognition for TSM native speakers. The results provide support for a dialect effect on TSM speakers' word recognition in accordance with prior exposure, native-ness and variant frequency.

2. Background

The central goal of this study is to understand how dialectal variation affects speakers' word recognition. The variation of interest is a phonemic difference in the vowel inventory in two TSM dialects, in which the difference signals a lexical contrast in one dialect but not in the other. Would words containing the target variations be recognized equally effectively, or would one form serve as a better prime over the other? Would the variations be processed differently according to the speakers' dialect, or would they be treated equally across different dialects? If one variant is treated as a better prime regardless of the speakers' dialect, what is the property of this form? This section provides a review of the previous literature on how non-contrastive elements, especially involving dialectal variations, may affect listeners' perception.

One major line of research in speech processing is concerned with the difference between phonemes and variants of the same phoneme (i.e. allophones in complementary distribution or free variation). Previous perceptual studies have generally shown that allophones form a single perceptual object in that allophones are perceived as the same phoneme and are perceived, or processed, as more similar than separate phonemes (e.g. Jaeger 1980; Sumner & Samuel 2005; Connine & Pinnow 2006; Ranbom & Connine 2007; Boomershine et al. 2008; Babel & Johnson 2010; Johnson & Babel 2010; Lu 2014). For example, in a category-formation experiment that asked English native listeners to identify whether the word they heard contained the category /k/, Jaeger (1980) showed that the listeners classified allophones of /k/ (e.g. aspirated, unaspirated, unreleased, palatalized, backed, labialized /k/) as instances of the same category, regardless of the spellings of the sound (i.e. *k*, *c, ch, q, x*).[1] This study suggests that, using an off-line identification task, allophones may be perceived as categorically similar.

The lack of effect of non-contrastive elements was also found in word recognition tasks (i.e. lexical decision – how quickly people classify stimuli as words or non-words). For example, Catalan contrasts segments such as /e/ vs. /ɛ/ while Spanish does not. In a priming paradigm in which a word with the target contrasts (e.g. /netə/) followed by either the repetition of the same word (e.g. /netə/) or its counterpart in a minimal pair (e.g. /nɛtə/) 8–20 items down the trial list, Pallier et al. (2001) showed that Spanish-dominant bilinguals exhibited a priming effect when the same words and words in a minimal pair were presented to them. Catalan-dominant bilinguals, however, showed a priming effect only when presented with the same words, suggesting an advantage of having contrastive elements in a speaker's dominant language in word recognition.

Similar effects were also found for free variations. For example, in a series of experiments using semantic priming, form priming and lexical decision tasks, Sumner & Samuel (2005) found that the target word *music* was immediately primed by the semantically related word *flute*, when *flute* was articulated with any of the three variants of final /t/: canonical fully aspirated [tʰ], coarticulated [ʔt̚] and glottalized [ʔ]. However, when the participants were presented with a contrastive phoneme /s/ instead of [t] (i.e. [flus]), no facilitation in classifying the target word *music* was observed. More significantly, when the primes and targets were presented in two different blocks, long-term priming was found only for the canonical fully aspirated [tʰ], rather than for the most frequent coarticulated [ʔt̚]. Based on the observed prototypicality effect and the lack of an observed frequency effect in the long-term priming paradigm, Sumner & Samuel (2005) concluded that exemplar representations are not stored in the long-term underlying representation, and that only elements stored in long-term memory serve as strong primes. These studies showed a robust effect of contrastive elements in sound categorization and word recognition and a limited effect of allophonic/free variations. Note, however, that the prototypicality or underlying effects found in the above studies usually coincide with the canonical representation of a certain sound (e.g. stressed onset for English [tʰ] and [kʰ]). In other words, one can argue that the lack of a non-contrastive effect may be due to meta-language processing, such as spelling or education.

However, several other studies have shown effects of allophonic knowledge on speech processing (e.g. Whalen et al. 1997; Mackenzie et al. 2014; Garcia &

1. In the category-formation task, the participants were first told that the "correct" category consisted of words containing the prototypical aspirated allophone [kʰ]. They were then tested to categorize other allophones of /k/.

Campos-Astorkiza 2015). In English, stops are produced with aspiration in the onset of stressed syllables, while the allophones, the unaspirated counterparts, occur in the onset of unstressed syllables. In a series of experiments, Whalen et al. (1997) asked English native speakers to perform perceptual preference and discrimination tasks on aspirated and unaspirated allophones of /p/ in the second syllable of disyllabic *non-words* with stress either on the first or on the second syllable to provide appropriate and inappropriate contexts (i.e. [p^h]/[p] as onset of stressed syllables and [p^h]/[p] as onset of unstressed syllables). The results showed that the aspirated [p^h] was preferred independent from the contexts and that the unaspirated [p] was more difficult to discriminate from the other allophone. However, when the same tasks were performed with *real words*, English native speakers did prefer the appropriate allophones. In other words, the results showed an allophonic effect only with real words, but not with non-words, suggesting that lexical knowledge may be required to induce non-contrastive, allophonic effects.

Along the same line, Garcia & Campos-Astorkiza (2015), using real words, induced a very similar allophonic effect on perception with Spanish /s/. In production, the Spanish intervocalic /s/ is voiced in a gradient fashion according to the following contexts from the most voiced to the least voiced: word-final context (e.g. *las alas*), word-initial (e.g. *la sopa*), and word medial (e.g. *casa*). In a similarity rating experiment in which Spanish speakers gave similarity judgments on pairs of /s/ sounds taken from these contexts, Garcia & Campos-Astorkiza (2015) showed that Spanish speakers rated sound pairs as significantly more different according to the contexts: word-medial, word-final, and then word-initial. The same results were replicated in a discrimination experiment, in which Spanish speakers gave more 'different' judgments to the sound pairs according to the same scale.

Beyond the realm of phonemes and free or allophonic variants, recent studies have also looked at dialectal differences. For example, the New York City dialect of English (NYC) is known to have *r* dropping in words such as *dark* [dɔək] and *sister* [sɪstə], while General American English (GA) maintains *r* production in these contexts. In a series of priming experiments employing both immediate as well as long-term priming paradigms, Sumner & Samuel (2009) showed that immediate priming exists between the "*r-ful*" and "*r-less*" forms for speakers with NYC exposure but not for speakers with GA exposure, suggesting an effect of prior experience to *r*-less forms (i.e. an exemplar effect); however, in long-term priming, the *r-less* forms did not prime as successfully as the *r-ful* forms, suggesting that both groups of speakers store only a single dialectal variant: the non-regional, *standard r*-forms. Note, however, that the standard forms (*r-ful* forms) coincide with the written form. Thus, it is possible that the processing for speakers of both dialects simply reflects the orthographic forms, in which *r* is always represented.

In a similar series of priming experiments, Chuang (2017) investigates immediate word recognition involving two kinds of "variations" in Taiwan Mandarin (TM): retroflexion vs. deretroflexion (i.e. dentals /ts, ts^h, s/ → retroflexes [tʂ, tʂ^h, ʂ] vs. retroflexes /tʂ, tʂ^h, ʂ/ → dentals [ts, ts^h, s]) and final nasals (e.g. /iŋ/→[in] vs. /in/→[iŋ]). Although dentals and retroflexes as well as /n/ and /ŋ/ codas contrast lexical meanings (e.g. /sa/ 'spread' vs. /ʂa/ 'kill' and /in/ 'sound' vs. /iŋ/ 'eagle'), variations are widely documented in TM (e.g. Chung 2006; Fon et al. 2011). Specifically, deretroflexion is found more frequently than retroflexion; /in/→[iŋ] is found predominantly in the northern dialect and enjoys a positive social connotation, while /iŋ/→[in] is found only in the southern dialect and is stigmatized. The results showed that while TM listeners were sensitive to the "standard" forms (i.e. the underlying form), the socially positive forms also facilitated word recognition. The frequency effect, compared to the social connotation effect, however, was found to be very limited in this study.

To summarize the discussion above, although previous studies have shown that non-contrastive elements do not enjoy the same status as contrastive elements, perceptual and processing differences across variants have been found in various degrees according to, but not limited to, the following factors: (1) the privileged status of the "standard" or "underlying" variants (e.g. the *r-ful* forms, the stressed onset for English [t^h] and [k^h]); (2) the word status: real words induce a non-contrastive effect more successfully than non-words; (3) degree of exposure to the variations (e.g. *r-less* dialect exposure, frequency); and (4) the type of tasks employed – an immediate word recognition paradigm induces a variation effect more successfully than long-term word recognition, off-line identification and discrimination tasks (e.g. Jaeger 1986; Sumner & Samuel 2005, 2009; Chuang 2017).

The present study builds on theses previous findings by conducting a long-term repetition priming task to investigate the processing of non-contrastive elements in two Taiwan Southern Min (TSM) dialects. TSM dialects serve as an ideal test case because, although these dialects are regionally different, there is no obvious "positive" vs. "negative" social connotation. Also, these dialects are spoken languages without a uniform writing system. While both dialects *can* be written either in Romanization or in Chinese characters, most speakers do not receive a literary education in this language. Even for those who are educated in the writing systems, these systems do not reflect dialectal differences. The advantage of examining the non-contrastive elements in TSM, as opposed to the languages investigated in previous studies, is the reduced effect from the underlying or orthographic influence from education. Furthermore, the current study employed a word recognition task that involved both words and non-words and recruited TSM speakers who were under 38 years of age, an age group that has maximal exposure to both dialects, to enhance the possibility of observing a dialectal effect (Hsu 2016). And finally, the current study used a long-term priming paradigm, as opposed to a short-term priming paradigm, to further test the findings of the previous research in which non-contrastive effects are mostly found at the level of short-term memory.

The following section presents background information on the two target dialects.

3. The target dialects in Taiwanese Southern Min

The origin, phonemic inventory, and geographical coverage of the dialects of Southern Min spoken in Taiwan are well documented (e.g. Chung 1996; Ang 2003; Chen 2010; Ang 2013; Hsu 2016). Chen (2010) and Hsu (2016) report on two emergent dialects, the "Five-Vowel" dialect and the "Six-Vowel" dialect (following Chen's terms), that show a phonemic contrast, as shown in Table 1.

Table 1. Two emergent Taiwan Southern Min dialects (Chen 2010)

Dialect	Five-Vowel dialect			Six-Vowel dialect		
Vowel inventory	i e	 a	u ɔ	i e	 ə a	u ɔ
Difference	**No /ɔ/-/ə/ contrast**			**/ɔ/-/ə/ contrast**		
Regions	Taipei (Northern Taiwan)			Tainan, Kaohsiung, Changhua (Central and southern Taiwan)		

The Five-Vowel dialect is argued to have evolved from the merging of the /o/ and /ɔ/ vowels in a previous six-vowel system /i, e, a, u, o, ɔ/ while the Six-Vowel dialect is argued to have evolved from a lip-spreading process of the /o/ to /ə/ (see a review in Chen 2009; 2010). These sound changes resulted in the fact that the Six-Vowel dialect contrasts /ɔ/ and /ə/ (e.g. /ɔ-a/ 'taro' vs. /ə-a/ 'oyster'), while the Five-Vowel dialect does not (e.g. *taro* and *oyster* are homophones: /ɔ-a/). More examples of this dialectal phonemic difference are listed below:

Table 2. Dialectal phonemic contrast

Five-Vowel	Six-Vowel	Gloss
[ɔ-a]	[ɔ-a] vs. [ə-a]	'taro' vs. 'oyster'
[hɔ]	[hɔ] vs. [hə]	'tiger' vs. 'good'
[t^hɔ]	[t^hɔ] vs. [t^hə]	'dirt' vs. 'peach'
[pɔ]	[pɔ] vs. [pə]	'clothes' vs. 'paper'
[k^hɔ]	[k^hɔ] vs. [k^hə]	'pants' vs. 'lesson'

A calculation of token frequency based on the *CCU Taiwanese Spoken Corpus* (Tsay & Myers 2013) indicates that syllables containing [ɔ] are twice as frequent compared to those containing [ə], as shown in the following table.

Table 3. Token frequency of [ɔ] and [ə] based on the *CCU Taiwanese Spoken Corpus*

	[ɔ]	[ə]
Token frequency	14,894	7,104

Influence from popular media and population movements for work and study have resulted in frequent contact between younger speakers of the two dialects. The age range in the young group reported in Hsu (2016) was from 21 to 33 years of age (born between 1983–1995).[2] It has been predicted that these two dialects will continue to influence other regionally adjacent dialects (Chen 2010; Hsu 2016). Given the lack of orthographic influence and obvious social connotations, these dialects provide a good opportunity to investigate the processing of non-contrastive information. And as noted in Sumner & Samuel (2009: 487), "the large majority of research on dialect variation has instead focused on the description of dialect, attitudes towards dialects, and the perception of vowel mergers across dialects", similar to the studies on Chinese dialects, of which "the examination of dialectal variation from a spoken word recognition standpoint has occurred relatively recently." The current research is one of the first studies that investigates word recognition in TSM dialects.

To present the TSM data, the writing convention for segmental and tonal contrasts in the Dictionary of Frequently-Used Taiwan Minnan compiled by the Ministry of Education is employed throughout the paper (http://twblg.dict.edu.tw/holodict_new/default.jsp). The target contrast is transcribed as *oo* [ɔ] vs. *o* [ə]. The following section details the experimental probe and the study's results.

4. Experiment

This study investigates the processing of two vowels [ɔ] and [ə] in Taiwan Southern Min by examining the lexical decision for words containing these vowels by speakers of two TSM dialects in which the two vowels are either contrastive (the Six-Vowel dialect) or not (the Five-Vowel dialect). A long-term repetition priming task was chosen as a probe to study the lexical representations of the two vowels. Previous

2. Chen (2010) investigated the production of TSM dialects from three groups: old, middle-age, young. The age range, however, was not specified.

studies employing a similar paradigm have shown that when participants are asked to perform lexical decisions on lists of stimuli in which some of the tokens appear twice, they respond more rapidly on their second encounter with the same word (e.g. Radeau et al. 1995; Pallier et al. 2001; Lee 2007; Sumner & Samuel 2007). Researchers therefore hypothesize that earlier activation of a lexical item facilitates later activation of the same item. The methodology of this study employed a word-recognition task similar to the one used in Pallier et al. (2001).

4.1 Methodology

4.1.1 *Participants*

Forty-three participants were recruited from National Chiao Tung University and received minor financial compensation for their participation. They first participated in a production pre-test that classified them into dialect groups based on their pronunciation of words containing [ɔ]-[ə] differences. The pre-test was a spontaneous guided introduction of the participants' family members (e.g. *I have an aunt. I don't have a brother*). A list of questions (e.g. *Do you have an aunt? Do you have brothers or sisters?*) was given in written Chinese. No spoken forms were given during the production pre-test. Kinship terms were used as a diagnostic of dialect group belongingness because these terms often reflect one's parental language identity. The kinship terms that contain the target vowels are *a-só* [a-sɔ]/[a-sə] 'sister-in-law', *î-pô* [i-pɔ]/[i-pə] 'great aunt', *sió-muē* [siɔ-mue]/[siə-mue] 'little sister', *a-koo* [a-kɔ] 'aunt', *a-tsóo* [a-tsɔ] 'great-grandfather/mother', and *tsa-bóo-kiánn* [tsa-bo-kiã] 'daughter'. If other words that contain the contrast were used in the participants' production, they were also taken into consideration (e.g. *bô* 'no'). If the participants produced the [ɔ]-[ə] contrast in these words, they were grouped into the Six-Vowel dialect, and if they did not produce the contrast, they were grouped into the Five-Vowel dialect. The judgments of dialect group belongingness were done separately by three trained phoneticians and jointly when there was a disagreement. The acoustic analyses of the production data will be presented in the next section. Although it has been pointed out in the previous literature that dialect production is not always representative of dialect perception and representation, the current study employed the standard definition in which a dialect is generally defined in terms of production (Sumner & Samuel 2009). Three of the participants displayed mixed dialect behavior. Their data were excluded from the data analysis. In total, data from 20 participants from each dialect group (Five-Vowel: 6M, 14F, ages 20–38; Six-Vowel: 8M, 12F, ages 22–37) were included in the final analysis. The regions where these participants were from are listed in the following table.

Table 4. Language background

	Five-Vowel	Six-Vowel
Age	20–38	22–37
Gender	6M, 14F	8M, 12F
Hometown	Taipei 5, Taoyuan 3, Hsinchu 3, Taichung 1, Changhua 1, Nantou 2, Yunlin 2, Chiayi 1, Tainan 1, Taitung 1	Taipei 2, Hsinchu 4, Taichung 1, Changhua 1 Chiayi 2, Tainan 6, Kaohsiung 1, Pingtong 2, Hualian 1,

Crucially, all the participants were Taiwan Southern Min speakers who were below the age of 38 at the time of the experiment, an age range that is comparable to the definition of the younger group in Hsu (2016).[3] In a questionnaire completed after the experiment (see Appendix A), these participants also reported that they had moved to different parts of Taiwan for work or study. Their backgrounds ensured that these participants had been exposed to different dialects prior to the recruitment. None of the participants reported any hearing deficiencies.

4.1.2 *Design and materials*

The stimuli consisted of 144 disyllabic words forming 72 minimal pairs. The minimal pairs differed either in the target contrast (*Specific contrast*: [ɔ] vs. [ə]; e.g. *ōo-á* 'taro' vs. *ô-á* 'oyster') or in a contrast shared by the two dialects, included for comparative purposes (*Common contrast*: [i] vs. [e]; e.g. *hî-á* 'fish' vs. *hê-á* 'shrimp'). These contrasts were embedded either in the first or second syllable of each disyllabic word (e.g. 1st syllable: ***ōo**-á* ~ ***ô**-á*; 2nd syllable: *tsit-kh**oo*** 'one dollar'~ *tsit-kh**o*** 'one subject'). In each contrast, three conditions were manipulated: (1) *Word condition*, in which both items in the minimal pair are words (e.g. *ōo-á* 'taro' vs. *ô-á* 'oyster'); (2) *[ə] only condition*, in which only the item containing [ə] is a word (e.g. **t**ōo**-lōo, tō-lōo* 'road'); and (3) *[ɔ] only condition*, in which only the item containing [ɔ] is a word (e.g. *ts**oo**-tshù* 'rent house', **ts**o**-tshù*). Note that these different conditions in the Specific contrast were developed on the basis of the Six-Vowel dialect, since the Five-Vowel dialect does not contrast [ɔ] and [ə]. In other words, in the Five-Vowel dialect, the minimal pairs in the Specific contrast were all homophones (e.g. [ɔ-a] 'taro/oyster') and the pairs in the *[ə] only* condition and the *[ɔ]*

3. A reviewer has pointed out that TSM speakers of this age group use Mandarin as the dominant language and hence they might not be fluent TSM speakers. The inclusion of the participants was based on three criteria. First, the included participants all have the ability of answering the questions and producing the lexical items included in the questions at a comfortable and fluent manner. Second, the included participants all reported that they used TSM at home with family members. Third, the included participants' self-rated TSM ability was all over 5 on a 1–7 scale (mean TSM ability: 5.71; mean Mandarin ability: 6.79).

only conditions were all produced with [ɔ]. These different conditions enabled us to investigate the effect of prior exposure because, although one form may not exist in one's dialect (e.g. *_t**ōo**-lōo_ in the Six-Vowel dialect), it is a legal and accepted form in the other dialect (e.g. Five-Vowel speakers produce 'road' as *_t**ōo**-lōo_). If prior exposure has an effect on speakers' word recognition, they should accept forms such as *_t**ōo**-lōo_ more readily than forms such as *_ts**o**-tshù_, a form that does not exist in either dialect.

Another 48 pseudo-word pairs (96 words) were included to serve as filler items to balance the word/non-word responses in the lexical decision task.[4] The manipulation of the stimuli is summarized in the following table and the complete list is in Appendix B.[5]

Table 5. Example stimuli

	Specific contrast **[ɔ] vs. [ə]**	**Common contrast** **[e] vs. [i]**
Position	1st syllable: _ōo-á_ 'taro' vs. _ô-á_ 'oyster' 2nd syllable: _tsi̍t-kh**oo**_ 'one dollar' vs. _tsi̍t-kh**o**_ 'one subject'	1st syllable: _hî-á_ 'shrimp' vs. _hê-á_ 'fish' 2nd syllable: _tuā-ke_ 'big chicken' vs. _tuā-ki_ 'big classifier'
Wordhood	Word: _ōo-á_ 'taro' vs. _ô-á_ 'oyster' [ə] only: *_tōo-lōo_ vs. _tō-lōo_ 'road' [e] only: _ké-sian_ 'pretentious' vs. *_kí-sian_	Word: _hî-á_ 'shrimp' vs. _hê-á_ 'fish' [ɔ] only: _tsoo-tshù_ 'rent house' vs. *_tso-tshù_ [i] only: *_pé-kàu_ vs. _pí-kàu_ 'compare'
	2 (Contrast) 2 (Position) 3 (Wordhood) 6 = 72 Exp pairs 48 pseudoword filler items, total 72+ 48 = 120 pairs	

A female phonetician who is a native speaker of the Six-Vowel dialect recorded multiple examples of the stimuli in a sound-treated booth, using a Marantz digital recorder PMD661 and AKG 220 microphone, at a sampling rate of 44,100 Hz. One instance of each stimulus was selected as the test item so that the tokens in a minimal pair were approximately matched on pitch and speaking rate. In order to

4. The filler pairs (48 pairs) were fewer than the experimental pairs (72 pairs) because some of the items in the experimental pairs were expected to have non-word responses, depending on the different Wordhood conditions across the two dialects.

5. I unfortunately was not able to balance the frequency across the experimental items due to the rarity of these minimal pairs and the lack of a large-scale TSM corpus. I have listed the word/morpheme frequency of the stimuli in Appendix B calculated based on _CCU Taiwanese Spoken Corpus_ (Tsay & Myers 2013).

control the intensity across tokens, the average intensity of each token was scaled to 65 dB, the rough average of the intensity of all the tokens, using Praat software (Boersma 2001).

Four stimuli lists were created so that each experimental word pair (e.g. *ōo-á* vs. *ô-á*) was presented in one of the following prime-target pairs: [ɔ] prime-[ɔ] target (e.g. *ōo-á* prime, *ōo-á* target), [ə] prime-[ə] target (e.g. *ô-á* prime, *ô-á* target), [ɔ] prime-[ə] target (e.g. *ōo-á* prime, *ô-á* target), and [ə] prime-[ɔ] target (e.g. *ô-á* prime, *ōo-á* target). In each list, the prime appeared first and was followed by its target 8-to-20 items down the list. The members of a given word pair appeared in the same positions in all four lists. The design of the experiment is shown in Table 6.

Table 6. Design of the experiment

Group	5-Vowel vs. 6-Vowel dialect
Prime-Target pairs	[ɔ] prime-[ɔ] target [ə] prime-[ə] target [ɔ] prime-[ə] target [ə] prime-[ɔ] target
Wordhood	Word vs. [ə] only vs. [ɔ] only

4.1.3 *Procedure*

Participants completed the experiment individually in a sound-treated booth, using a computer connected to a serial response box with keys labeled 'yes' or 'no' in Chinese characters. The four stimulus lists were presented binaurally over headphones at a comfortable listening level, using E Prime software (Schneider et al. 2002) and the participants were randomly assigned to one of the lists. Participants were presented with written instructions on the computer screen in Chinese and were asked to make a lexical decision (to judge if the stimulus was a word or not) by pressing 'yes' or 'no' on the response box as soon as they were sure, to secure an accurate response time. Each participant completed a practice session with 8 trials before the experiment and had time to ask questions. The experiment lasted approximately 20 minutes.

4.2 Results

To ensure that the experimental stimuli and the dialect group belongingness were appropriate, the stimuli and the participants' production data were analyzed and presented in §4.2.1. Word-recognition results are presented subsequently in §4.2.2. The subsequent analyses were conducted in R (R Core Team, 2017), and the visual presentation of the data was carried out using ggplot2 (Wickham 2009) and effects (Fox 2003) packages.

4.2.1 *Acoustic analyses*

In this section, two sets of analyses were done, one on the validity of the experimental items, and one on the appropriateness of the dialect grouping.

To see if there was a significant acoustic difference between the two target vowels of the experimental items (36 pairs forming 72 items), the F1 and F2 were measured at the midpoint of the vowels using Praat (Boersma 2001). The mean values of the F1 and F2 of the two vowels are shown in Table 7 with standard deviations in parentheses. The following scatter plot shows F1 at the midpoint as a function of F2 at the midpoint of the target vowels for the 72 items.

Table 7. Mean values of F1 and F2 (in Hz) for the target vowels in the experimental stimuli with standard deviations in parentheses

	F1	F2
[ə]	522.59 (43.56)	1597.24 (86.32)
[ɔ]	576.81 (38.54)	939.34 (59.42)

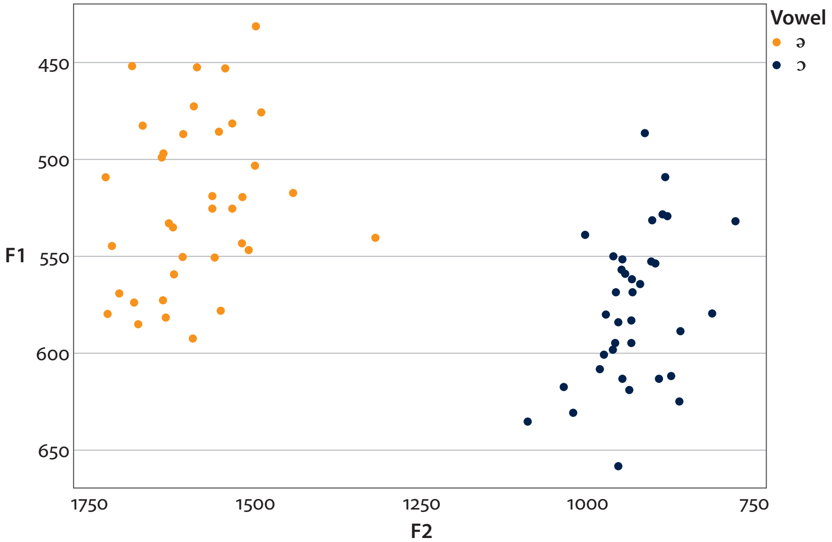

Figure 1. F1 at the midpoint as a function of F2 at the midpoint of the target vowels in the stimuli

Two paired-samples *t*-test were conducted to compare the F1 and F2 taken from the midpoints of the target vowels. The results showed that there was a significant difference in both the F1 ($t(35) = 7.51, p < .001$) and F2 ($t(35) = -41.615, p < .001$) of the two vowels.

Another paired-samples *t*-test was conducted to compare the duration of the experimental stimuli that contained the two target vowels. The results did not yield a significant difference ($t(35) = -1.5137$, $p = 0.14$), indicating that the durations of the two vowels were comparable. The durations of the two vowels are thus omitted from the following acoustic analyses. The descriptive results are shown in Table 8.

Table 8. Mean duration of the experimental stimuli containing the two vowels with standard deviation in parentheses

[ə]	649.36 (73.98)
[ɔ]	658.51 (76.56)

To ensure the appropriateness of the dialect grouping, the same measurements were carried out on the participants' production of the target vowels. The descriptive data are shown in Table 9 and the corresponding scatter plot is in Figure 2.

Table 9. Mean values of F1 and F2 (in Hz) for the target vowels in the participants' production pretest with standard deviations in parentheses

		F1	F2
Six-Vowel	[ə]	595.13 (122.94)	1250.65 (215.48)
	[ɔ]	625.92 (77.45)	989.33 (151.55)
Five-Vowel	[ɔ]	596.70 (124.55)	1003.26 (220.14)

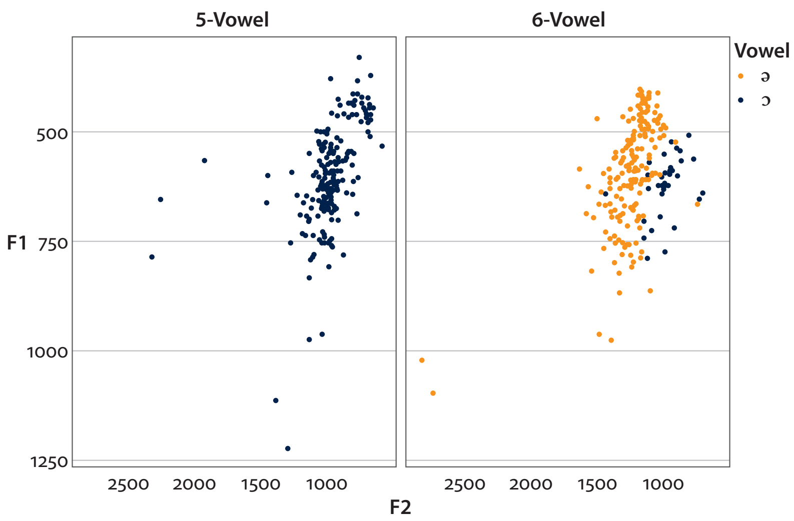

Figure 2. F1 at the midpoint as a function of F2 at the midpoint of the target vowels in the participants' production pretest

We can see from the scatter plot that the [ɔ]s produced by the Six-Vowel and Five-Vowel groups show a high degree of overlap, while the [ə]s produced by the Six-Vowel group form a distinct cluster. Due to the unbalanced number of items of the two vowels produced by each participant in the guided spontaneous speech, the results were analyzed with a linear mixed-effects regression model using the lme4 package of R (Bates et al. 2015). Mixed-effect regression modeling has a number of advantages over analysis of variance (ANOVA) in that it allows for the inclusion of participant and item random effects as well as both continuous and factorial variables in one statistical model. The models were fitted with the F1 and F2 midpoint measurements (in Hz) as the dependent variables. The fixed effects included in the models were Dialect Group (i.e. Five-Vowel vs. Six-Vowel) and Vowel (i.e. [ə] and [ɔ]), and the random effects were Participant and Item.

Two models were fitted to compare the two vowels for the Six-Vowel group to ensure that the two vowels produced by the participants were significantly different. The results showed that the two vowels were significantly different for F2 ($\beta = 299.75$, SE = 68.95, $p < .001$) but not for F1 ($\beta = -3.19$, SE = 26.21, $p = 0.91$). These results indicate that the difference between the vowels lies in backness, signaled mainly by F2 differences (Table 11), and not in vowel height, signaled mainly by F1 differences (Table 10).

Table 10. Six-Vowel F1 ~ Vowel + (1|Participant) + (1|Item)

	Estimate	Standard error	t value	Pr (> |*z*|)
(Intercept)	603.172	31.389	19.216	<.001
Vowel	−3.197	26.214	2.730	0.911

Table 11. Six-Vowel F2 ~ Vowel + (1|Participant) + (1|Item)

	Estimate	Standard error	t value	Pr (> |*z*|)
(Intercept)	981.967	65.868	14.908	<.001
Vowel	299.750	68.954	4.347	<.001***

These results ensured that the participants grouped as Six-Vowel speakers did have the vowel distinction in their production. Note that the two vowels are distinct in both dimensions in the experimental stimuli but distinct only in the dimension of F2 in the participants' production. This could be due to the fact that the experimental stimuli were produced with more care. The following scatter plots show the acoustic information of vowels in the stimuli and those produced by the participants of the two dialect groups for comparative purposes.

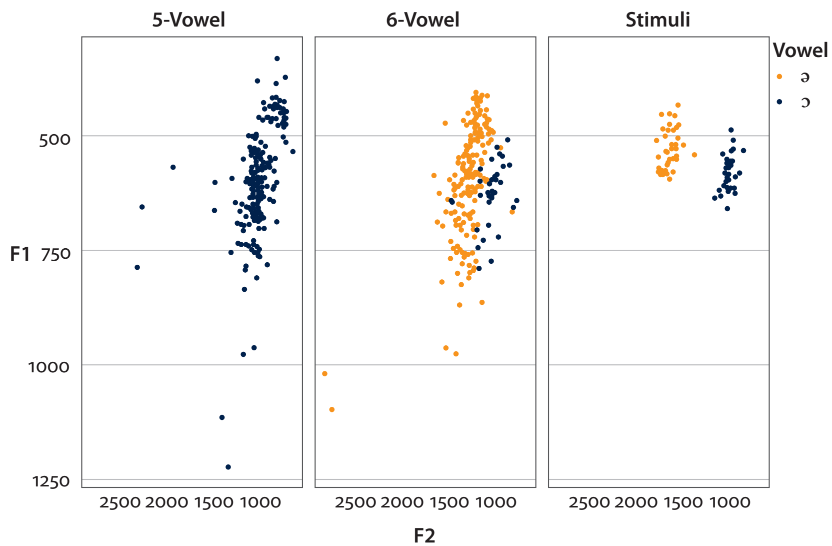

Figure 3. F1 at the midpoint as a function of F2 at the midpoint of the target vowels paneled by the dialect group and experimental stimuli

To see if the [ɔ]s produced by the Five-Vowel group form one category and are comparable to the [ɔ] category and distinct from the [ə] category produced by the Six-Vowel group, models were fitted to compare the [ɔ]s produced by the two groups. The results showed that the [ɔ]s produced by the two groups were not significantly different in the dimensions of F1 (β = 8.03, SE = 35.71, p = .82; Table 12) and F2 (β = 12.85, SE = 61.41, p = .83; Table 13).

Table 12. F1 ~ Group + (1|Participant) + (1|Item)

	Estimate	Standard error	t value	Pr (> |z|)
(Intercept)	614.84	24.12	25.48	<.001
Vowel [ɔ]	8.03	35.71	0.22	=.82

Table 13. F2 ~ Group + (1|Participant) + (1|Item)

	Estimate	Standard error	t value	Pr (> |z|)
(Intercept)	1025.20	48.59	21.10	<.001
Vowel [ɔ]	12.85	61.41	0.21	=.83

Models were further fitted to compare the [ɔ]s produced by the Five-Vowel group and the [ə]s produced by the Six-Vowel group to see if they were significantly different. The results showed that the two vowels were different in the dimension of F2 ($\beta = -233.10$, SE = 44.12, $p < .001$; Table 15) but not of F1 (Table 14), suggesting that the two vowels were significantly different in backness.

Table 14. F1 ~ Vowel + (1|Participant) + (1|Item)

	Estimate	Standard error	t value	Pr (> |z|)
(Intercept)	601.77	25.18	23.897	<.001
Vowel [ɔ]	12.24	34.91	0.35	=.73

Table 15. F2 ~ Vowel + (1|Participant) + (1|Item)

	Estimate	Standard error	t value	Pr (> |z|)
(Intercept)	1262.23	44.15	28.59	<.001
Vowel [ɔ]	−233.10	44.12	−5.28	<.001***

The above acoustic analyses excluded the possibilities that the [ə] and [ɔ] are not produced equivalently by speakers of the two dialects, resulting in a different phonemic categorization, and in turn, a different perception, and that the vowel [ɔ] in the Five-Vowel dialect may exist on a pronunciation continuum somewhere between the Six-Vowel [ə] and [ɔ], again causing different phonemic categorization.

4.2.2 *Lexical decision results*

The central goal of this study is to investigate how speakers of the Five-Vowel group and the Six-Vowel group process words that contain the target vowels [ə] and [ɔ] in a word recognition task. Reaction times more than 2.5 standard deviations (sd = 543.01 ms) from the mean (mean = 1244.23 ms) were discarded (<4%).[6] Data were further screened with the error rates to the non-critical items (i.e. items in the Common contrast). Four participants (1M, 1F from the Five-Vowel group and 2F from the Six-Vowel group) with error rates greater than 40% were excluded from the analysis.[7] Mean reaction times for the four different prime-target pairs in the

6. Reaction times were measured at the onset of the stimulus presentation.

7. Note that the cut-off point (40%) for exclusion based on the error rate was higher than that in Sumner & Samuel (2009) (10%), a paper with a similar design that looks at the same issue. The response times reported in this study were in general longer than those in Sumner & Samuel (2009) as well (all means below 1000 ms). This may be due to the fact that the participants recruited in this study were also native speakers of Mandarin Chinese and that Mandarin was the dominant language of these participants. It is likely that they processed TSM in a slower fashion and with higher error rate.

Specific contrast of the two dialect groups are provided in Table 16. A corresponding break-down of the error rates is given in Table 17.

Table 16. Mean reaction times in millisecond in the Specific contrast

	Five-Vowel		Six-Vowel	
	Prime	**Target**	**Prime**	**Target**
[ɔ] prime-[ɔ] target	1008.95	898.55	999.67	865.70
[ə] prime-[ə] target	1111.98	985.56	1109.04	1023.36
[ɔ] prime-[ə] target	1194.86	1119.05	1135.15	1114.23
[ə] prime-[ɔ] target	1122.53	930.00	1016.83	967.86

Table 17. Mean error rates in the Specific contrast

	Five-Vowel		Six-Vowel	
	Prime	**Target**	**Prime**	**Target**
[ɔ] prime-[ɔ] target	0.27	0.29	0.22	0.25
[ə] prime-[ə] target	0.24	0.20	0.17	0.17
[ɔ] prime-[ə] target	0.25	0.38	0.17	0.31
[ə] prime-[ɔ] target	0.18	0.30	0.09	0.28

Repetition effects were then calculated by measuring the decrease in lexical-decision response times between the first and second occurrences of an item ([ɔ] prime-[ɔ] target, [ə] prime-[ə] target) or between the occurrence of an item and its counterpart in a minimal pair ([ɔ] prime-[ə] target, [ə] prime-[ɔ] target). If one of the items in a prime-target pair was excluded in the screening process described above, the given pair was discarded. A positive number signals a repetition effect, while a negative number signals an inhibition effect. A number that is close to zero signals no priming.

In fitting models to this dataset, *Repetition Effect* is considered the dependent variable. *Dialect Group, Prime-Target Pairs*, and *Wordhood* are the experimental variables of interest. Additionally, other variables were considered in each model. *Participant* and *Item* were included as random effects. Control variables were also included, allowing item property (*Position*), participant property (*Gender*) and experiment property (*Trial Order*) to influence the response variable. A description of each variable is listed in Table 18.

A model was first fitted with all available predictor variables and was then compared to models with variables removed one by one. In this procedure, *Position* and *Gender* were eliminated for their lack of significant improvement to the model

Table 18. Variables considered for analysis

Group	5-Vowel vs. 6-Vowel dialect
Prime-Target pairs	[ɔ] prime-[ɔ] target [ə] prime-[ə] target [ɔ] prime-[ə] target [ə] prime-[ɔ] target
Wordhood	Word vs. [ə] only vs. [ɔ] only
Position	Target vowels as 1st or 2nd syllable in the disyllabic stimuli
Gender	Gender of the participants
Trial Order	Presentation order of stimuli
Participant	Participant ID (random effect)
Item	Experimental items (random effect)

as indicated by likelihood ratio testing.[8] *Trial Order*, however, did significantly contribute to the goodness-of-fit to the data ($\chi^2 = 9134$, $p < .01$), as shown in Figure 4.

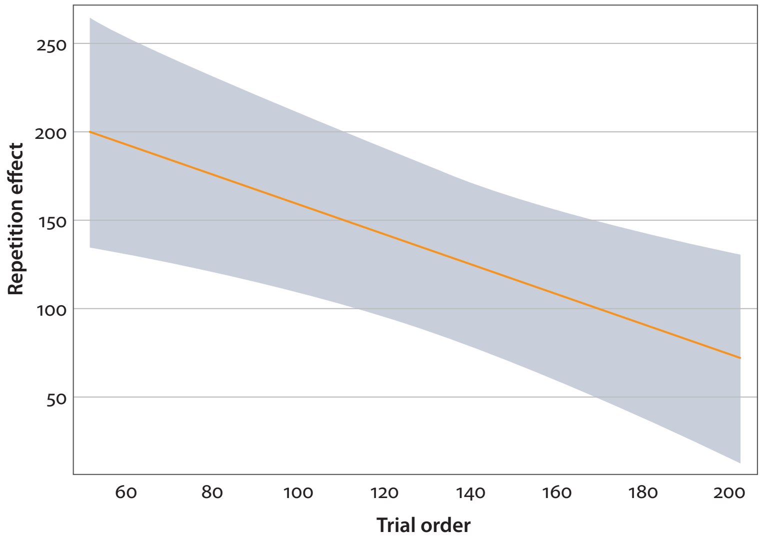

Figure 4. Effect of trial order

This showed that there was a negative correlation between repetition effect and trial order, suggesting that listeners' lexical decision responses were facilitated with more practice, an effect that has been previously observed (Keuleers et al. 2010). In the following analysis, only the main effects of the experimental variables (*Dialect Group, Prime-Target Pair, Wordhood*) and their interactions were included. The

8. Chi-squared statistics and *p*-values were used to compare two models (Baayen et al. 2008).

repetition effects for the two dialect groups across the four *Prime-Target* pairs paneled by *Wordhood* is shown in Figure 5.

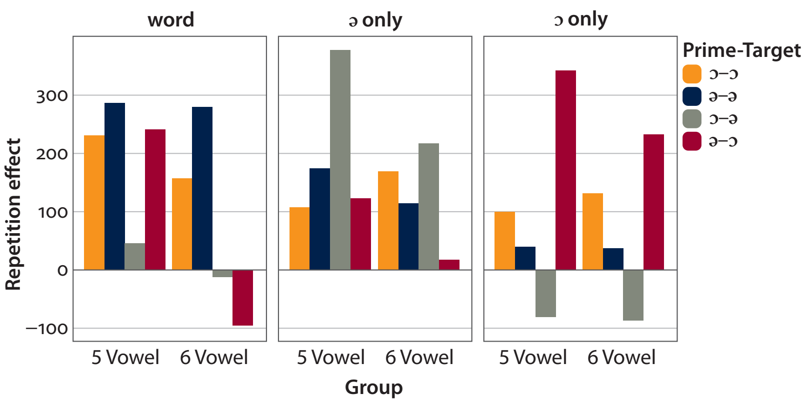

Figure 5. Repetition effects

To further interpret the results, two sets of analyses using mixed-effects regression are presented: the first containing data from both groups and the second performed on each *Prime-Target* pair. The first analysis revealed a *Dialect Group* main effect ($p < .05$), suggesting that the participants in the Five-Vowel group had an overall larger repetition effect than those in the Six-Vowel group. Two interactions were also found: one of *Dialect Group*Prime-Target Pair* ($p < .05$), and the other of *Prime-Target Pair*Wordhood* ($p < .001$).[9]

Table 19. Repetition effect ~ Dialect Group*Prime-Target Pair*Wordhood + (1|Participant) + (1|Item)

	F. value	Pr(>F)
Dialect Group	4.76	<.05*
Wordhood	1.88	0.16
Prime-Target Pair	1.29	0.29
Dialect Group : Wordhood	1.34	0.26
Dialect Group : Prime-Target Pair	3.04	<.05*
Wordhood : Prime-Target Pair	6.88	<.001***
Dialect Group : Wordhood : Prime-Target	1.01	0.41

The interactions suggest that words containing [ə] and [ɔ] presented in different *Prime-Target* pairs were recognized with different speeds based on the speaker's dialectal background and the *Wordhood* status of the stimulus.

9. The *p*-values of the main effects and interactions calculated based on Satterthwaite's approximations were obtained using the lmerTest package of R (Kuznetsova et al. 2016).

Further mixed-effect analyses were fitted for each of the *Prime-Target* pairs. The lack of main effects and interactions in the "ɔ-ɔ" and "ə-ə" pairs showed that when the prime and target were identical, repetition effects were not different between the two dialect groups or across different *Wordhood* manipulations. These results were expected because earlier activation of a lexical item facilitates later activation of the same item (items in *Word* condition, [ə]-items in *[ə] only* condition, and [ɔ]-items in *[ɔ] only* condition) and non-word processing can also be facilitated by an earlier presentation of a similar non-word ([ɔ]-items in *[ə] only* condition and [ə]-items in *[ɔ] only* condition) (Sumner & Samuel 2007).

Table 20. "ɔ-ɔ" Prime-Target pair

	F. value	Pr(>F)
Dialect Group	0.01	0.91
Wordhood	0.49	0.62
Dialect Group : Wordhood	0.91	0.41

Table 21. "ə-ə" Prime-Target pair

	F. value	Pr(>F)
Dialect Group	0.16	0.68
Wordhood	3.4	0.07
Dialect Group : Wordhood	0.12	0.87

Models fitted on the data when the prime and target were minimal pairs, however, revealed interesting findings. The results for the minimal pairs were replotted in Figure 6 for comparative purposes.

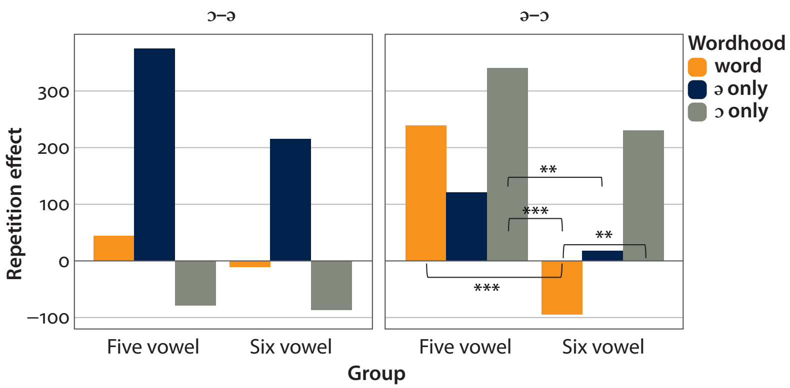

Figure 6. "ɔ-ə" and "ə-ɔ" Prime-Target pairs

First, in “ɔ-ə” *Prime-Target* pairs (left panel of Figure 6), a *Wordhood* main effect was found, as shown in Table 22. Post-hoc analyses showed that it was driven by the difference between the *[ə] only* condition and the other conditions (*Word* condition: $p < .05$; *[ɔ] only* condition: $p < .01$).[10] The lack of repetition effects for the *Word* condition was expected since words containing the two vowels are contrastive in this condition. In other words, presenting an [ɔ]-word followed by an [ə]-word counterpart did not yield facilitation because the two words simply activated different lexical items. Second, the same lack of priming effects was also found for the *[ɔ] only* condition. That is, no facilitation was found when presenting an [ɔ]-word followed by its [ə]-non-word counterpart. Facilitation, however, was found for the *[ə] only* condition. That is, an [ɔ]-non-word primed its [ə] word counterpart. Implications of these results will be discussed in the next section.

Table 22. “ɔ-ə” Prime-Target pair

	F. value	Pr(>F)
Dialect Group	1.69	0.2
Wordhood	11.11	<.01**
Dialect Group : Wordhood	0.71	0.49

The results for the “ə-ɔ” *Prime-Target* pairs revealed even more interesting results, as shown in the right panel of Figure 6. Main effects of *Dialect Group* and *Wordhood* were found as well as their interaction, as shown in Table 23. Post-hoc analyses showed that the interaction was driven by the following pairs: (1) Five-Vowel vs. Six-Vowel in *Word* condition; (2) Six-Vowel *Word* condition vs. Six-Vowel *[ɔ] only* condition; (3) Six-Vowel *Word* condition vs. Five-Vowel *[ɔ]-only* condition; and (4) Six-Vowel *[ə] only* condition vs. Five-Vowel *[ɔ] only* condition.

Table 23. “ə-ɔ” Prime-Target pair

	F. value	Pr(>F)
Dialect Group	16.64	<.001***
Wordhood	10.51	<.001***
Dialect Group : Wordhood	3.23	<.05*

The first comparison indicated that facilitation was found for the Five-Vowel group but not for the Six-Vowel group for the *Word* condition ($p < .001$), contrary to what

10. A Tukey procedure was used for post-hoc tests to control the family-wise error rate over the entire set of pair-wise comparisons. The post-hoc tests were conducted using the lsmeans package of R (Lenth 2016). The subsequent post-hoc analyses were run similarly.

was found for the “ɔ-ə” *Prime-Target* pair in which fewer or no repetition effects were found for the *Word* condition. That is, independent of dialect group, [ɔ]-words did not prime their [ə]-word counterparts while [ə]-words served as effective primes for their [ɔ]-word counterparts for the speakers in the Five-Vowel group. The second and third comparisons indicated that when the [ə]-item in a minimal pair was non-word, it primed a later presented [ɔ]-word counterpart independent of dialectal background. This contrasts sharply with the results when [ɔ]-word was presented before its [ə]-non-word counterparts (i.e. *Word* and *[ɔ] only* conditions in “ɔ-ə” *Prime-Target* pair). The fourth comparison indicated that an [ə]-non-word primed its [ɔ]-word counterpart in the Five-Vowel group, but an [ə]-word did not prime its [ɔ]-non-word counterpart in the Six-Vowel group.

5. Discussion and conclusion

A long-term repetition priming experiment investigating word recognition involving two vowels, [ə] and [ɔ], by speakers of Taiwanese Southern Min revealed that the two vowels were processed the same yet differently according to dialectal background and word status. Several interesting patterns were found:

a. In the *Word* condition, no repetition effect was observed when [ɔ]-words were presented prior to their [ə]-word counterparts across two dialects; [ə]-words did serve as effective primes when they were presented prior to their [ɔ]-word counterparts, but only for the Five-Vowel group.
b. Independent of dialect group, [ɔ]-non-words facilitated their [ə]-word counterparts, but [ɔ] words did not facilitate their [ə]-non-word counterparts.
c. Independent of dialect group, [ə]-non-words facilitated their [ɔ]-word counterparts.

The first pattern confirmed the contrastiveness of the two vowels in the Six-Vowel participants' word recognition, in that hearing minimal pairs of words containing the target vowels activated different lexical items and thus did not exhibit a priming relationship. However, this pattern indicated that the Five-Vowel speakers exhibited a mixed behavior in that the two vowels seem to have been treated contrastively when [ɔ]-words were presented first, but were treated as variants when [ə]-words were presented first. In other words, the results showed distinctive patterns from the two dialect groups, confirming the appropriateness of the dialect grouping based on the speakers' production. The phonemic status of the vowels does not appear to explain this asymmetrical behavior. If the two vowels are simply variants to the Five-Vowel speakers, we should have observed repetition effects in both “ɔ-ə” and “ə-ɔ” *Prime-Target* pairs; if the two vowels are contrastive, as in the Six-Vowel dialect, we should have observed no repetition effects in both *Prime-Target* orders. Frequency does not seem to explain this pattern either. As mentioned earlier, syllables containing [ɔ] are twice as frequent as those containing [ə]. If frequency plays a role, we would have expected [ɔ]-words to serve as more effective primes and would have found a stronger repetition effect for the "ɔ-ə" *Prime-Target* pair. The fact that the repetition effect was found only for the "ə-ɔ" but not "ɔ-ə" *Prime-Target* pair calls for another explanation. Why would [ə]-words serve as better primes than [ɔ]-word counterparts but not vice versa? One possibility for this behavior may be attributed to the "non-nativeness" of the [ə] vowel in that the Five-Vowel participants might have taken more time to process words containing [ə], and thus when [ə]-words were presented later, word recognition was delayed, thus resulting in the lack of a repetition effect. This possibility was supported by subsequent analyses on raw lexical decision times to words containing [ə] and [ɔ] by speakers of the two dialect groups. Two mixed-effect regression models with lexical decision times as dependent variable, *Vowel* (i.e. [ə] and [ɔ]) and *Appearance* (Prime vs. Target) as independent variables, and *Participant* and *Item* as random effects, showed that although speakers from both dialect groups took more time in making lexical decisions for primes than for targets, only the Five-Vowel speakers consistently took more time in response to words containing [ə].

Table 24. Five-Vowel group response time ~ Vowel*Appearance+ (1|Participant) + (1|Item)

	F. value	Pr(>F)
Vowel	8.63	<.01**
Appearance	71.68	<.01***
Vowel : Appearance	0.20	0.64

Table 25. Six-Vowel group response time ~ Vowel*Appearance + (1|Participant) + (1|Item)

	F. value	Pr(>F)
Vowel	0.98	0.98
Appearance	37.49	<.001***
Vowel : Appearance	2.02	0.15

The dialect-specific behavior, however, was canceled out when considering *Wordhood*, as in the second and third patterns. Uniform results across dialect groups were found in which [ɔ]-non-words primed their [ə]-word counterparts but [ɔ]-words did not prime their [ə]-non-word counterparts. This may be attributed to the fact that these [ɔ]-non-words are legal productions of the [ə]-words in the Five-Vowel group. That is, although **too-lo* [tɔ-lɔ] is not a word in the Six-Vowel dialect, it is the actual production of the same word *to-lo* [tə-lɔ] 'road' in the Five-Vowel dialect. In other words, prior dialectal exposure may have enabled the speakers to interprete these words with their [ə]-word counterparts (the [ɔ]s in these non-words were produced as [ə]s in the Six-Vowel dialect) and hence resulted in repetition effects. However, when [ɔ]-words were presented first and their [ə]-non-word counterparts later, prior exposure did not provide these speakers with any [ə]-non-word input (e.g. *[tsə-tsu] is not a word in both dialects), and hence resulted in no repetition effects.

As for the third pattern, in which [ə]-non-words facilitated their [ɔ]-word counterparts, it could be attributed to the fact that word recognition may also be facilitated after a similar sounding non-word (e.g. Connine et al. 1993; Sumner & Samuel 2007). Previous studies have reported that a lexical item can be activated by similar-sounding non-words that differ in one or two linguistic features (e.g. voicing in ***p**attern* vs. *****b**attern*), while non-words that deviate by more than three linguistic features show no priming effect (e.g. voicing, place of articulation and manner in ***p**attern* vs. *****r**attern*). I therefore attribute this to the phonetic similarity between the targets and their [ə]-non-word counterparts.

The above interpretation of the results was based on the assumption that these variants are stored in long-term memory and thus have an effect on speakers' word recognition in an experimental paradigm such as the one used in the current study. These findings lend support to an alternative view presented in previous studies in which an immediate word recognition paradigm induces a variation effect more successfully than long-term word recognition, off-line identification and discrimination tasks. In fact, the results of the current study showed a limited effect of contrastive elements in that minimal pairs containing [ɔ] and [ə] were not simply processed as the same by the Five-Vowel group and were not simply processed as different by the Six-Vowel speakers. The current study, setting aside the privileged status of the "canonical" or "underlying" variants stemming from orthography or education, provides possible evidence for non-contrastive effects in long-term word recognition.

However, one may be concerned that the repetition effects found in the Five-Vowel group may not have been caused by prior exposure to the Six-Vowel dialect, but were instead caused by the fact that these listeners only have [ɔ] in their inventory. Under this hypothesis, when Five-Vowel speakers heard words containing [ə], they rejected them, categorizing them as non-words, and accepted only the tokens containing [ɔ] as words. In this case, the repetition effect may have been caused simply by a prior presentation of a similar-sounding word or non-word, as has been found in previous studies (e.g. Connine et al. 1993; Connine et al. 1997; Sumner & Samuel 2007). If this interpretation is correct, we would expect Five-Vowel speakers' lexical decisions to be different when minimal pairs are presented, regardless of different *Wordhood* conditions. In other words, when presented with minimal pairs, speakers of this dialect should accept [ɔ]-words while rejecting [ə]-words. To evaluate this possibility, lexical decision correspondences between the two occurrences of the same item ("ɔ-ɔ" and "ə-ə" *Prime-Target* pairs) and of the minimal pairs ("ə-ɔ" and "ɔ-ə" *Prime-Target* pairs) were calculated. If lexical responses to presentations of the same item and of minimal pairs were cross-consistent (both accepted as words or both rejected as non-words), the value 1 was given; if the lexical responses to the presentations were not consistent (e.g. one was accepted as a word, while the other was not), the value 0 was given. A number closer to 1 signals that, for a given participant, either the same lexical representation was activated by the presentations of the same item and minimal pair, or that both pairs were perceived as non-words. A number closer to 0 signals that the participant accepted one of the pairs as a word and rejected the other. If the Five-Vowel listeners rejected all words containing [ə] as non-words, we should find low word-correspondence scores when presented with minimal pairs across different *Wordhood* conditions in this dialect. The results are shown below.

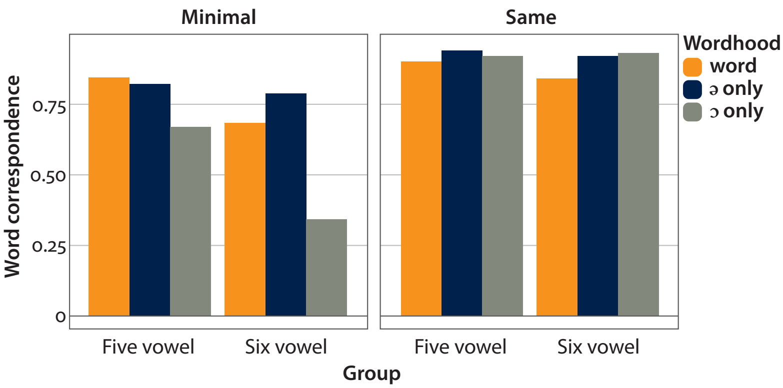

Figure 7. Word correspondences between two occurrences of the same item or of minimal pairs in percentages

The figure shows that the word correspondences in lexical decisions made by speakers of the two dialects were high in the right panel (around 90%), indicating that they consistently judged presentations of the same items twice as either words or non-words. The correspondences when minimal pairs were presented, however, were lower (around 70%). Mixed-effect logistic regression models were run with *Word Correspondence* as the dependent variable, *Dialect Group* (Five-Vowel vs. Six-Vowel), *Pair* (same vs. minimal pair) and *Wordhood* (*Word, [ə] only, [ɔ] only*) as fixed variables, and Participant and Stimulus as random variables. The results revealed significant *Dialect Group* ($\chi^2 = 11.74, p < .001$) and *Pair* ($\chi^2 = 14.24, p < .001$) main effects. Crucially, we also found significant *Pair*Dialect Group* ($\chi^2 = 4.44, p < .05$), *Pair*Wordhood* ($\chi^2 = 6.26, p < .05$) and *Pair*Wordhood*Dialect Group* interactions ($\chi^2 = 19.37, p < .01$).[11] In other words, participants' word correspondences differed according to different the *Wordhood* conditions and to the dialect group that they belonged to. The most interesting results here lie in the word correspondences observed when minimal pairs were presented. Post-hoc tests revealed that the Six-Vowel group had higher word correspondences in the *Word* and *[ə] only* conditions compared to the *[ɔ] only* condition (both pairs $p < .05$). The Five-Vowel group showed the same trend but without statistic significance. This suggests that they were more likely to judge words containing [ə] as non-words in the *[ɔ] only* condition. This is expected if we take prior exposure into consideration, since the tokens containing [ə] in this condition were non-words in both dialects. However, the higher word correspondences in the *[ə] only* condition suggest that tokens containing [ə] were judged as words at the same rate by both the Five-Vowel speakers and the Six-Vowel speakers. In other words, speakers from both dialects processed the minimal pairs similarly in the *[ə] only* condition but not in the *[ɔ] only* condition. If dialectal information is not stored in long-term memory, such an asymmetrical pattern should not arise.

Interestingly, a group effect was also found in the minimal pairs in the *[ɔ] only* condition ($p < .001$), suggesting that Six-Vowel speakers rejected [ə]-words more readily than Five-Vowel speakers did. This finding supports the interpretation that a frequency effect occurs in long-term word recognition. That is, since the occurrences of these words are more frequent in the Six-Vowel speakers' exemplars and less frequent in the Five-Vowel speakers' exemplars, the rejection of words containing [ə] was more consistent by the Six-Vowel speakers than by the Five-Vowel speakers.

To summarize, the current study investigated word recognition involving non-contrastive, dialectal information – the two vowels, [ə] and [ɔ], are contrastive in Six-Vowel TSM but non-contrastive (homophonous) in Five-Vowel TSM. The results of a long-term repetition priming experiment showed that speakers of the two dialects recognized words containing [ə] and [ɔ] in the same yet different fashions according to their dialectal background and to prior exposure to the other

11. To obtain main effect *p*-values, a full model was run with all three independent variables. This model was then compared with models with each of the variables removed one at a time in a likelihood ratio test using the anova () function in R. The *p*-values for the interactions were obtained similarly by comparing models without the interaction term (e.g. Word Correspondence ~ Dialect Group + Wordhood + (1|Participant) + (1|Item)) and with the interaction term (e.g. Word Correspondence ~ Dialect Group*Wordhood + (1|Participant) + (1|Item)).

dialect. Specifically, no repetition effects were found for speakers across dialect groups when words containing [ɔ] were presented prior to their [ə]-word counterparts. This suggests that speakers from both groups may have treated the two vowels differently. That is, even when one's native dialect does not contrast the two vowels, these non-contrastive elements may also affect one's word recognition in long-term priming paradigms. However, a dialectal difference was observed when [ə]-words were presented prior to their [ɔ]-word counterparts. This is attributed to the non-nativeness of [ə]-words to the Five-Vowel speakers so that words containing [ə] were recognized at a slower pace. Furthermore, contrary to what has been found in previous studies, in which one variant consistently serves as a better prime than the other form, the current study found that these variants may have a different facilitating effect according to the word status that they are involved in. This is attributed to prior exposure to the other dialect – when the items containing the target vowels are words in the other dialect, the items were recognized more readily than those that are not words. A frequency effect was also found in word correspondence rates, which showed that Six-Vowel speakers more consistently rejected words containing [ə] compared to Five-Vowel speakers.

I conclude with two remarks and a possible direction for future research. First, previous studies have failed to elicit non-contrastive effects with long-term priming paradigms. For example, Sumner & Samuel (2005) showed that, among the various pronunciations of final /t/, only the canonical fully aspirated [t^h] served as an effective prime in a long-term priming paradigm. And as was also reviewed earlier, Pallier et al. (2001), using a similar paradigm to investigate the priming relationship between sounds that are contrastive in Catalan but not in Spanish, showed that a priming relationship exists when minimal pairs involving these sounds were presented 8–20 words apart, as well as when the same item was presented twice for Spanish-dominant bilinguals but not for Catalan-dominant bilinguals. In an even more relevant study in which non-contrastive, dialectal variation was investigated, Sumner & Samuel (2009) showed that, independent of dialectal background, only the *r ful* form facilitated a later presented *r-less* form, but not the other way around in long-term priming despite the existence of immediate priming between the *r-ful* and *r-less* forms for speakers with NYC English exposure. The question is then why a non-contrastive effect was found in the current study in a long-term priming paradigm. One possible difference between this study and the previous studies is that there is no influence from orthography and education to strengthen the "canonical" or "standard" forms. Such forms may have caused a stronger priming effect from the canonical/standard forms in the previous studies. Without such forms, TSM for example, there might be a need to store different variants in long-term memory. Another possibility is that the TSM speakers recruited in this study were also native speakers of Mandarin, in which the two target vowels, [ə] and [ɔ], are semi-contrastive (e.g. /təŋ/ 'climb' vs. /tɔŋ/ 'east').[12] These two vowels are contrastive after non-labial onset and before velar nasal coda in Mandarin. The prestige status of Mandarin may have caused a higher sensitivity to the contrastive nature of the two sounds for the Five-Vowel speakers. However, this factor alone cannot explain the asymmetrical pattern displayed between the speakers of the two dialects.

Second, one may wonder why the repetition effects elicited in the current study were so much stronger than those in the previous studies (over 200 ms in the current study vs. below 100 ms in the previous studies). This may be attributed to the fact that Taiwanese, along with other Sinitic languages, is an isolating language in which monomorphemic polysyllabic words are extremely rare. According to a corpus calculation in Hsieh (2006), there are less than 50 items found. Hsieh (2006: 2) noted that "the one-morpheme-per-syllable tendency is arguably robust." As a result, most of the stimuli in minimal pairs compiled in the current research inevitably involved semantic relatedness (e.g. *kò-lîn* 'personal' vs. *kòo-lîn* 'a deceased person'). There were only a few that were not related semantically (e.g. *pòo-kíng* 'background display' vs. *pò-kíng* 'call the police'). This may have caused the larger priming effect in the current study.

A logical follow up would be to investigate word recognition differences among speakers with different degrees of exposure as well as differences between immediate priming and long-term priming. These will be left for future research.

Acknowledgements

I would like to thank James Myers, Janice Fon, Samuel Wang, Shu-chuan Chen, and two anonymous reviewers for their valuable comments and suggestions. I would also like to thank my research assistants Jia-Wen Wei, Yi-Hsuan Kuo, Dai-ying Lin, and Wei-Fang Hsieh for their help in running the experiment. Any remaining errors are mine. This work was supported by Ministry of Science and Technology of Taiwan Grant (MOST103-2628-H-009-001 and MOST104-2410-H-009-031) to the author.

12. I thank James Myers for bringing this factor to my attention.

Appendix A. 測後問卷

請問您使用那些語言(包括您的母語)? 您何時學習這些語言? 以1-7的評分，標示您每一個語言的能力; 1為能力低7為使用能力類似母語使用者。

1. 語言名稱及學習年齡 語言能力:

語言名稱及學習年齡	低						高
___________________	1	2	3	4	5	6	7
___________________	1	2	3	4	5	6	7
___________________	1	2	3	4	5	6	7
___________________	1	2	3	4	5	6	7
___________________	1	2	3	4	5	6	7
___________________	1	2	3	4	5	6	7

2. 請問您的種族 (請於合適處複選):
   _____ 原住民 _____ 亞裔 _____ 太平洋島族
   _____ 非裔 _____ 白種 _____ 西班牙裔
   _____ 南亞 _____ 其他 (請標示: _________________)
3. 請問您日常使用語言是?
   ____
   ____
4. 請問您曾住過的城市、國家，以及居住時之年齡及長度。從出生地開始列，直到最近的居住地。
   城市，國家 居住年齡
   例: 柏林，德國 例: 出生到八歲
   ____
   ____
   ____
   ____
5. 請勾選您的最高學歷。
   無正規教育 __________
   小學/基礎教育 __________
   中學/國中 __________
   高中職 __________
   技職 __________
   大學 __________
   碩士 __________
   法醫博士 __________
   博士 __________
6. 請問您的性別是? ____
7. 請問您的職業是? ____
8. 請問令尊的職業是? ______________________________________
9. 請問令堂的職業是? ______________________________________
10. 請問令尊使用的語言是? ___________________________________
11. 請問令堂使用的語言是? ___________________________________

12. 請問您今年貴庚? ______________________________
13. 請問您有聽力及語言上的障礙嗎? 如果有,請標示。
____
14. 請問您覺得這個實驗的測試內容為何?

謝謝您的參與!

Appendix B. Stimulus list

Token frequencies are listed for the disyllabic words or for each morpheme in the words if the former is not available from *CCU Taiwanese Spoken Corpus* (Tsay & Myers 2013).

	oo [ɔ]	Gloss	Token frequency	*o* [ə]	Gloss	Token frequency	Wordhood
1	kòo-lîn	'old friend'	64/1911	kò-lîn	'indivudial'	31	Word
2	póo-phín	'nutritious food'	6	pó-phín	'treasure'	13/325	Word
3	ōo-á	'taro'	6	ô-á	'oyster'	11	Word
4	kôo-á	'glue'	9	ko-á	'cake'	7	Word
5	pòo-kíng	'display'	37/18	pò-kíng	'call the police'	6	Word
6	hōo-miâ	'account name'	14/366	hō-miâ	'name' v.	519/366	Word
7	tsi̍t-hōo	'one account'	2805/14	tsi̍t-hō	'number one'	2805/519	Word
8	tsi̍t-khoo	'one dollar'	2805/77	tsi̍t-kho	'one subject'	2805/249	Word
9	tuā-koo	'aunt'	451/16	tuā-ko	'big brother'	451/5	Word
10	tuā-too	'big city'	451/323	tuā-to	'big knife'	451/6	Word
11	âng-thôo	'red soil'	337/53	âng-thô	'red pitch'	337/45	Word
12	tsin-khóo	'bitter'	1762/121	tsin-khó	'condense'	1762/0	Word
13	*soo-á			so-á	'rope'	0	[ə] only
14	*poo-lê			po-lê	'glass'	0	[ə] only
15	*thóo-tsînn			thó-tsînn	'ask for money'	8/201	[ə] only
16	*tōo-lōo			tō-lōo	'road'	6/34	[ə] only
17	*tsóo-á			tsó-á	'date'	0	[ə] only
18	*hóo-lâng			hó-lâng	'nice person'	2486/1911	[ə] only
19	*pua̍h-tóo			pua̍h-tó	'fall down'	0/23	[ə] only
20	*hué-koo			hué-ko	'hot pot'	36/0	[ə] only
21	*huân-lóo			huân-ló	'worry'	7/0	[ə] only
22	*tsuí-kóo			tsuí-kó	'fruit'	26	[ə] only
23	*kî-tóo			kî-tó	'pray'	2	[ə] only
24	*hóo-lâng			hó-lâng	'nice guy'	2	[ə] only
25	lôo-hué	'stove'	0/36	*lô-hué			[ɔ] only
26	lóo-bī	'braised dishes'	1/64	*ló-bī			[ɔ] only
27	koo-tsiánn	'invite someone'	0/0	*ko-tsiánn			[ɔ] only
28	too-tshī	'city'	323/2	*to-tshī			[ɔ] only

	oo [ɔ]	Gloss	Token frequency	*o* [ə]	Gloss	Token frequency	Wordhood
29	tsoo-tshù	'rent house'	3/90	*tso-tshù			[ɔ] only
30	tshoo-sio̍k	'vulgar'	11/21	*tso-sio̍k			[ɔ] only
31	a-tsóo	'grandfather'	23	*a-tsó			[ɔ] only
32	tsiàu-kòo	'look after'	64	*tsiàu-kò			[ɔ] only
33	piān-sóo	'bathroom'	25	*piān-só			[ɔ] only
34	té-khòo	'shorts'	10/7	*té-khò			[ɔ] only
35	ē-poo	'afternoon'	6	*ē-po			[ɔ] only
36	tshng-khòo	'warehouse'	0/16	*tshng-khò			[ɔ] only

References

Ang, Ui-jin. 2003. *Yinbian de dongji yu fangxiang: Zhang-Quan jingzheng yu Taiwan putung-qiang de xingcheng* [The motivation and direction of sound change: On the competition of Minnan dialects Chang-chou and Chuan-chou, and the emergence of general Taiwanese]. Hsinchu: National Tsing Hua University. (Doctoral dissertation.)

Ang, Ui-jin. 2013. Taiwan de yuzhong fenbu yu fenqu. [The distribution and regionalization of varieties in Taiwan]. *Language and Linguistics* 14(2). 315–369.

Baayen, R. & Davidson, D. & Bates, D. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Babel, Molly & Johnson, Keith. 2010. Accessing psycho-acoustic perception and language-specific perception with speech sounds. *Laboratory Phonology* 1. 179–205. https://doi.org/10.1515/labphon.2010.009

Babel, Molly & Russell, Jamie. 2015. Expectations and speech intelligibility. *Journal of Acoustical Society of America* 137. 2823–2833. https://doi.org/10.1121/1.4919317

Bates, Douglas & Mächler, Martin & Bolker, Benjamin & Walker, Steve. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5. 341–345.

Boomershine, Amanda & Hall, Kathleen Currie & Hume, Elizabeth & Johnson, Keith. 2008. The impact of allophony versus contrast on speech perception. In Avery, Peter & Dresher, Elan & Rice, Keren (eds.), *Contrast in Phonology*, 145–171. Berlin: Mouton de Gruyter.

Chen, Shu-chuan. 2009. Taiwan Minnanyu yuanyin xitong ji yin-yang rushengdiao de bianyi yu bianhua-Taiwan Minnanyu de zibiao diaocha fenxi [The vowel system change and the Yin-/Yang-entering tonal variations in Taiwanese Hokkien]. *Journal of Taiwanese Languages and Literature* 3. 157–178.

Chen, Shu-chuan. 2010. Taiwan Minnanyu xinxing de yuyin bianyi-Taibeishi, Zhanghuashi ji Tainanshi yuanyin xitong yu yangru yuandiao de diaocha fenxi [New sound variation in Taiwan Southern Min: Vowel systems and the lower register entering tone in Taipei, Changhua, and Tainan]. *Language and Linguistics* 11(2). 425–468.

Chomsky, Noam & Halle, Morris. 1968. *The sound pattern of English*. Cambridge: Harper & Row.

Chuang, Yu-Ying. 2017. *The effect of phonetic variation on word recognition in Taiwan Mandarin*. Taipei: National Taiwan University. (Doctoral dissertation.)

Chung, Karen Steffen. 2006. Hypercorrection in Taiwan Mandarin. *Journal of Asian Pacific Communication* 16. 197–214. https://doi.org/10.1075/japc.16.2.04chu

Chung, Raung-fu. 1996. *The segmental phonology of Southern Min in Taiwan*. Taipei: Crane Publishing Co., Ltd..

Clements, G. N. 2003. Feature economy in sound systems. *Phonology* 20. 287–333. https://doi.org/10.1017/S095267570400003X

Connine, Cynthia M. & Blasko, Dawn G. & Titione, Debra. 1993. Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language* 32. 193–210. https://doi.org/10.1006/jmla.1993.1011

Connine, Cynthia M. & Pinnow, Eleni. 2006. Phonological variation in spoken word recognition: Episodes and abstractions. *The Linguistic Review* 23. 235–245. https://doi.org/10.1515/TLR.2006.009

Connine, Cynthia M. & Titone, Debra & Deelman, Thomas & Blasko, Dawn. 1997. Similarity mapping in spoken word recognition. *Journal of Memory and Language* 37. 463–480. https://doi.org/10.1006/jmla.1997.2535

Fon, Janice & Hung, Jui-mei & Huang, Yi-Hsuan & Hsu, Hui-ju. 2011. Dialectal variations on syllable-final nasal mergers in Taiwan Mandarin. *Language and Linguistics* 12(2). 273–311.

Fox, John. 2003. Effect displays in R for generalised linear models. *Journal of Statistical Software* 8(15). 1–27. https://doi.org/10.18637/jss.v008.i15

Garcia, Christina & Campos-Astorkiza, Rebeka. 2015. Allophonic perception in Spanish: The case of intervocalic /s/ voicing. (Paper presented at the Phonetics and Phonology in Europe 2015, Cambridge, 29–30 June 2015.)

Hsieh, Feng-fan. 2006. High infidelity: The non-mapping of Japanese accent onto Taiwanese tone. *MIT Working Papers in Linguistics* 52. 1–27.

Hsu, Hui-ju. 2016. Biandong zhong de Taiyu: Taiyu /o/ yinsu sanzhong zhuyao duyin de xianzhuang fenxi. [The dynamics of Taiwanese – the analysis of the Taiwanese phoneme /o/]. *Dongwu Zhongwen Xuebao* [Soochow Journal of Chinese Studies] 31. 303–328.

Jaeger, Jeri J. 1980. Testing the psychological reality of phonemes. *Language and Speech* 23. 233–253. https://doi.org/10.1177/002383098002300302

Jaeger, Jeri J. 1986. Concept formation as a tool for linguistic research. In Ohala, John J. & Jaeger, Jeri J. (eds.), *Experimental Phonology*, 211–237. Orlando: Academic Press.

Johnson, Keith & Babel, Molly. 2010. On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics* 38. 127–136. https://doi.org/10.1016/j.wocn.2009.11.001

Kenstowicz, Michael & Kisseberth, Charles. 1979. *Generative phonology: Description and theory*. New York: Academic Press.

Keuleers, Emmanuel & Diependaele, Kevin & Brysbaert, Marc. 2010. Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono-and disyllabic words and nonwords. *Frontier in Psychology* 1. No. 174. 1–15.

Kuznetsova, Alexandra & Brockhoff, Per Bruun & Christensen, Rune Haubo Bojesen. 2016. lmerTest: Test in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26.

Lee, Chao-Yang. 2007. Does horse activate mother? Processing lexical tone in form priming. *Language and Speech* 50. 101–123. https://doi.org/10.1177/00238309070500010501

Lenth, Russell V. 2016. Least-squares means: The R package lsmeans, *Journal of Statistical Software* 69(1). 1–33. https://doi.org/10.18637/jss.v069.i01

Lu, Yu-An. 2014. Mandarin fricatives redux: The psychological reality of phonological representations. *Journal of East Asian Linguistics* 23. 43–69. https://doi.org/10.1007/s10831-013-9111-5

Mackenzie, Sara & Olson, Erin & Clayards, Meghan & Wagner, Michael. 2014. Allophonic variation in English /l/: Production, perception, and segmentation. (Paper presented at the 2014 Canadian Linguistic Association conference, St. Catharines, 24–26 May 2014.)

Pallier, Christophe & Colomé, Angels & Sebastián-Gallés, Núria. 2001. The influence of native-language phonology on lexical access: Examplar-based versus abstract lexical entries. *Psychological Science* 12. 445–449. https://doi.org/10.1111/1467-9280.00383

Patterson, D. & Connine, Cynthia M. 2001. Variant frequency in flap production: A corpus analysis of variant frequency in American English flap production. *Phonetica* 58. 254–275. https://doi.org/10.1159/000046178

Radeau, Monique & Morais, Jose & Segui, Juan. 1995. Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance* 21(6). 1297–1311.

Ranbom, Larissa J. & Cynthia M. Connine. 2007. Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language* 57. 273–298. https://doi.org/10.1016/j.jml.2007.04.001

Schneider, Walter. & Eschman, Amy. & Zuccolotto, Anthony. 2002. *E-Prime user's guide*. Pittsburgh: Psychology Software Tools.

Sumner, Meghan & Samuel, Arthur G. 2005. Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language* 52. 322–338. https://doi.org/10.1016/j.jml.2004.11.004

Sumner, Meghan & Samuel, Arthur G. 2007. Lexical inhibition and sublexical facilitation are surprisingly long lasting. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(4). 769–790.

Sumner, Meghan & Samuel, Arthur G. 2009. The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language* 60. 487–501. https://doi.org/10.1016/j.jml.2009.01.001

Tsay, Jane & Myers, James. 2013. *CCU Taiwanese Spoken Corpus*. (http://lngproc.ccu.edu.tw/SouthernMinCorpus/) (Accessed 2019-05-03)

Whalen, D. H. & Best, Catherin T. & Irwin, Julia R. 1997. Lexical effects in the perception and production of American English /p/ allophones. *Journal of Phonetics* 25. 501–528. https://doi.org/10.1006/jpho.1997.0058

Wickham, Hadley. 2009. *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag. https://doi.org/10.1007/978-0-387-98141-3

Author's address

Yu-An Lu
Department of Foreign Languages and Literatures
National Chiao Tung University
HC403, 4F, Humanities Bldg. 3
1001 University Road, Hsinchu 30010
Taiwan
yuanlu@nctu.edu.tw

Publication history

Date received: 21 December 2016
Date accepted: 29 December 2017