ORIGINAL ARTICLE

Open Access

CrossMark

On the dialectal variations of voiced sibilant /dz/ in Taiwan Min young speakers

Yu-Ying Chuang[1,2] and Janice Fon[1,3*]

* Correspondence: jfon@ntu.edu.tw
[1]Graduate Institute of Linguistics, National Taiwan University, 1 Roosevelt Rd. Sec. 4, 317 Lexue Bldg., Taipei 10617, Taiwan, ROC
[3]Graduate Institute of Brain and Mind Sciences, National Taiwan University, Taipei, Taiwan, ROC
Full list of author information is available at the end of the article

Abstract

This study examined the realization of /dz/ in three dialects of Taiwan Min among young speakers based on a paragraph-reading task. Results showed there were five robust realization categories of /dz/, including the commonly reported dental sibilants, velar obstruents, and liquids and the rarely mentioned dental nonsibilants and retroflexes. Liquids and retroflexes were more likely to occur before a rounded segment, while dental sibilants, dental nonsibilants, and velar obstruents were more likely to occur before an unrounded segment. For the latter environment, there was also a correlation between structural complexity and realization of dental sibilants, velar obstruents, and liquids. The three dialects examined demonstrated different attitudes towards the realization of /dz/. 泉 *Chôan* was the most receptive to the new sounds of liquids, dental nonsibilants, and retroflexes, while Mix was the most conservative in preserving the old forms of dental sibilants and velar obstruents. 漳 *Chiang* was somewhere in between. It maintained a robust dental sibilant category like Mix yet welcomed new variants of dental nonsibilants and retroflexes like *Chôan*. Talkers were consistent in their /dz/ realization choices and intra-talker variability was low. This study thus demonstrated that talkers' realization of a variable sound is a complex product of linguistic structure, speakers' background, and idiosyncratic preference.

Keywords: Taiwan Min; Voiced sibilant /dz/; Dialectal variation; Intra-speaker variability

1 Background

While voiceless sibilants are rather common among languages around the world, voiced ones are not. The UCLA Phonological Segment Inventory Database (Maddieson 1984) showed that out of the 451 languages investigated, as many as 414 incorporate voiceless sibilants in their inventory, while only 232 include voiced sibilants, which is a 92 vs. 51% ratio[a]. Among the 232 languages that have voiced sibilants, 147 have sibilant fricatives while 149 have sibilant affricates. A more recent P-base database (Mielke 2007) indicated that out of the 553 languages studied, 493 have voiceless sibilants, while only 304 have voiced ones, which is an 89 vs. 55% ratio[b]. Among the 304 languages with voiced sibilants, 206 have sibilant fricatives while 202 have sibilant affricates. Although the two databases do not include the same set of languages, and might have even adopted different criteria for phoneme decisions, the distributions are fairly similar, i.e., voiced sibilants are far less likely to occur than voiceless ones, and voiced sibilant fricatives and voiced sibilant affricates are equally likely to appear.

The much lower frequency of occurrence for voiced sibilants seems to be due to articulatory constraints. Ohala (1983) claimed that more stringent aerodynamic requirements are imposed by the articulation of voiced fricatives compared to their voiceless counterparts. In order to maintain voicing, the intraoral pressure has to be low relative to the subglottal pressure to facilitate vocal fold vibration. However, to create friction noise, the intraoral pressure has to be high relative to the atmospheric pressure outside the oral cavity to create high air velocity through the oral stricture. It is thus articulatorily demanding to meet both requirements simultaneously. In order to maintain sustained voicing, the intraoral pressure might not be high enough to create the necessary friction noise, resulting in voiced approximants. On the other hand, to maintain fricative noise, the intraoral pressure might not be low enough to facilitate voicing throughout the segment, resulting in voiceless fricatives. A successful voiced fricative thus requires a delicate balance between meeting the aerodynamic needs for both voicing and frication.

The situation is even more complicated with affricates, a composite articulation of a plosive and a fricative. For plosives, the feasibility of voicing is largely dependent on its duration (Ohala 1983). Like fricatives, voicing requires low intraoral pressure relative to the subglottal pressure. If the closure duration is relatively short, the intraoral pressure buildup can be well within the required range for voicing. However, if the closure duration is long, the intraoral pressure buildup can rise to a point at which there is no more transglottal pressure difference, and consequently voicing ceases. For the fricative portion of an affricate, the voicing dilemma is even more severe than a regular standalone fricative (Żygis 2008). As the plosive portion of an affricate is released into the fricative portion, the high pressure buildup at the end of the closure would tend to create an intraoral pressure level that is too high to maintain voicing, resulting in a (partially) devoiced affricate. This might be why for languages that incorporate a voiced affricate, there is a tendency for them to suppress intraoral pressure buildup by having shorter closure duration (Żygis et al. 2012) and to use voiced affricates only marginally (e.g., restricted to loan words or proper names only) (Żygis 2008).

Voicing contrasts might also be disadvantageous for sibilants from the acoustic and auditory perspectives. Balise and Diehl (1994) measured the RMS amplitude of frication noise in English sibilants and nonsibilants and found that voicing effectively reduces the frication noise of sibilants so that they become less distinguishable from their non-sibilant counterparts. Perceptual experiments also verified this by showing that voiced fricatives are more difficult to identify than voiceless ones, and the impact is more severe on the perception of sibilants than nonsibilants (Balise and Diehl 1994; Miller and Nicely 1955; Singh and Black 1966; Wang and Bilger 1973).

The stringent aerodynamics in voiced sibilants and their less distinctive acoustic and perceptual features seem to make this sound type quite variable. For example, /z/ is likely to be devoiced in voiceless environments in English (Smith 1997) and in word-initial positions in Dutch (Gussenhoven and Bremmer 1983). Similarly, the devoicing rates of word-final /z/and /ʒ/ in Portuguese can be as high as 93% (Jesus and Shadle 2003). The situation is even more variable with sibilant affricates. For instance, /dz/ is realized as [z] in certain morphological endings in Upper Sorbian (Schuster-Šewc 1999). In Gulf Arabic, /dz/ is realized as [gʲ] or [j] in certain contexts (Holes 1990). Similarly, /dʒ/ in Haitian Creole is realized as [j] in word-final positions (Tinelli 1981).

In other words, both voiced sibilant fricatives and affricates tend to undergo some phonological processes to accommodate the production and perception demands imposed by the sound category.

It is precisely from this perspective that this study finds the realization of the voiced sibilant in Taiwan Min worth investigating. Taiwan Min has its origin in Southern Min, a Chinese language spoken in southeastern provinces of China, including southern Zhejiang, southern Fujian, eastern Guangdong, and Hainan (Norman 1988)[c]. It is the second largest language spoken in Taiwan, and more than 70% of the population claim having at least some passive knowledge of Taiwan Min (Huang 黃宣範 1993). According to Ang 洪惟仁 (2013), there are three major dialects of Southern Min spoken in Taiwan, 漳 *Chiang*, 泉 *Chôan*, and Mix, with uneven distributions[d]. *Chiang* speakers are dominant in the west central inland areas, while *Chôan* speakers mainly reside in northern Taiwan and the west central coastal strip[e]. Speakers in southwestern Taiwan predominantly speak a variety of Min with a more balanced mixture of *Chiang* and *Chôan*, and is thus appropriately called the Mix dialect (Ang 洪惟仁 2005; Li 李仲民 2009).

Taiwan Min includes only one voiced sibilant in the system, traditionally symbolized as *j* in the Church Romanization system (Cheng and Cheng Xie 鄭良偉, 鄭謝淑娟 1977). Ang 洪惟仁 (2003) argued that there are two realization variants, [z] and [dz], and they show a dialectal split, with the *Chiang* and the Mix dialect being more inclined to realize it as [z], and *Chôan* being more inclined to realize it as [dz]. Both dialects demonstrate a palatalized realization of [z]/[dz] when the sound precedes an unrounded high vowel /i/ or glide /j/. In order to proceed in a more succinct manner, this study followed the literature tradition and used /dz/ to refer to this sound when its cross-dialectal phonemic status was intended [e.g., Ang 洪惟仁 (2012); Wang 王薈雯 (2014); Yao 姚榮松 (1988)]. However, readers should take note that variable pronunciations ranging from [z] to [dz] were intended by this symbol.

The variation of Taiwan Min /dz/ does not end here. A new variant [l] has evolved and become the dominant realization of /dz/ in the *Chôan* dialect (Ang 洪惟仁 2003, 2012; Ang and Chang 洪惟仁, 張素蓉 2008; Hung 洪慧鈺 2007; Thoo 涂文欽 2009; Wang 王薈雯 2014). For young speakers, [l] is almost the sole realization for /dz/ in both read (Ang 洪惟仁 2003, 2012; Ang and Chang 洪惟仁, 張素蓉 2008; Hung 洪慧鈺 2007; Thoo 涂文欽 2009) and spontaneous speech (Wang 王薈雯 2014). Even for speakers older than 60, many showed a predominant [l] realization rate of over 75%, pushing the original [z]/[dz] realization to a minor status (Ang 洪惟仁 2003, 2012; Hung 洪慧鈺 2007). In other words, the [z]/[dz] realization is not only rapidly disappearing among younger speakers but is also losing ground in the older generations. Even though there is still a positive correlation between [z]/[dz] and speaker age, [l] has become the new favorite across all age groups.

The situation is more complex in the *Chiang* dialect. In addition to [z]/[dz] and [l], there is also an additional [ɡ] variant (Ang 洪惟仁 2003, 2012), which is likely due to language contact with Hakka, the third largest language spoken in Taiwan (Huang 黃宣範 1993), as its occurrence coincides nicely with areas where the 鶴佬客 *Hoh-ló-kheh*s, or the "Min-Hakkas" (the Hakka people who cannot speak Hakka but speak Min natively), reside (Ang 洪惟仁 2012; Chuang et al. 莊雅雯等 2009). Ang's 洪惟仁 (2012) account for the influence of Hakka on the [ɡ] realization of Min /dz/ is rather circuitous. As /dz/ is not a sound indigenous to Hakka, it might have posed some difficulties

for Hakka learners of Min. However, many /dz/-initial words in Min are cognately related to [ɲ]-initial words in Hakka, and an association between /dz/ and [ɲ] might have thus been established. As [ɲ] is an allophonic variant of /ŋ/ in Hakka that only occurs before /i j/, Hakka learners of Min showed free variation of [ɲ] and [ŋ] for /dz/-initial words in Min due to negative transfer. Since [ŋ], being an allophone of /ɡ/ in Min, is a sound in Min phonology while [ɲ] is not, the former eventually won out and [ŋ] (and also its allophonic partner [ɡ]) became one of the common variants in realizing Min /dz/ for Hakka learners of Min. In other words, the [ɡ] realization of /dz/ in Min is created based on a calculated (albeit unconscious) compromise between the two systems by Hakka learners of Min. This tendency persists even after some Hakka people stop speaking Hakka and acquire Min natively without the direct negative interference from Hakka, and has gradually evolved into the [ɡ] realization in the *Chiang* dialect, the geographical areas of which are contiguous to the Hakka regions. Following the Hakka phonotactic constraints for [ɲ], the [ɡ] variant is only realized before high unrounded vowels or glides, while [z]/[dz] and [l] do not show such restrictions and could occur before both rounded and unrounded vowels/glides.
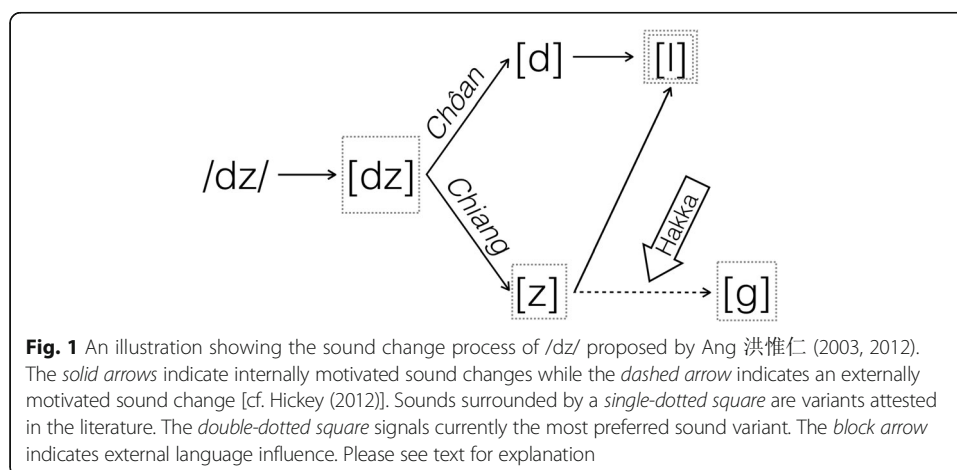
The three-way competition among [z]/[dz], [l], and [ɡ] before an unrounded high vowel/glide is age-dependent, and is especially intense between the former two, with [ɡ] being a rather minor form, the realization rate of which hovers only around 10% at most for all age groups (Ang 洪惟仁 2003, 2012). As in the *Chôan* dialect, the realization rate of [l] is in general negatively correlated with speaker age, while that of [z]/[dz] is positively correlated. For speakers above age 40, [z]/[dz] is still a strong match for [l] in certain syllable types, but it is losing ground among the younger speakers. However, unlike their counterparts in the *Chôan* dialect, *Chiang* speakers in general do not demonstrate a complete conversion from [z]/[dz] to [l]. Even for teenagers, the [z]/[dz] realization rates still maintain above 20%. On the other hand, the correlation between the use of [ɡ] and speaker age is not as straightforward. It is used predominantly among speakers between ages 20 and 60, but is not as preferred by speakers outside this age range, demonstrating a polynomial relationship with speaker age.

Most studies agreed that the three variants observed in the *Chiang* dialect are also found in the Mix dialect. [z]/[dz] is the dominant variant for old speakers, while [ɡ] and [l] are rather marginal (Ang 洪惟仁 1997; Chen 陳淑娟 1995; Lin 林珠彩 1995). For middle-aged speakers, [ɡ] and [l] are on the rise, while [z]/[dz] is gradually diminishing. Some claimed that [ɡ] is the dominant realization for this age group, with [l] lagging behind (Chen 陳淑娟 1995; Lin 林珠彩 1995), while others found that the two are in strong competition with each other (Khng 康韶真 2014). Some speakers use both [ɡ] and [l] freely before high front unrounded vowels/glides (Chen 陳淑娟 1995), while others show complementary distribution (Lin 林珠彩 1995), with [l] occurring only before high back rounded vowels/glides and [ɡ] occurring only before high front unrounded ones. For young speakers, [l] becomes more robust and eventually wins the competition, with a realization rate of over 85% (Khng 康韶真 2014). The realization rate for [z]/[dz] hovers around 10% only and that for [ɡ] is virtually nonexistent. Chen 陳雅玲 (2010, 2012) claimed that the distribution of the three variants is not only age-dependent but also location-specific. For example, [z]/[dz] is still fairly robust among Kaohsiung young speakers in the Xiaogang district but is rarely observed in the Zuoying district even for old speakers. Similarly, although [ɡ] is a trait said to be of

middle-aged talkers, it is hardly used by talkers of any age group from the Dalinpu community. However, Zuoying residents are fervent supporters of this variant, and nearly 20% of the tokens from young speakers and more than 35% of the tokens from old speakers were realized as such. In general, the distribution of /dz/ realization in the Mix dialect seems to parallel nicely with that in the *Chiang* dialect, with [z]/[dz] being the predominant variant for old speakers and [l] becoming the predominant variant for young speakers. However, unlike their *Chiang* counterparts, Mix speakers are more inclined to use [ɡ], and their middle-aged speakers even consider this as the dominant variant. For young speakers, [ɡ] is diminishing in a location-dependent fashion, with its use ranging from fairly robust to virtually nonexistent.

According to Ang 洪惟仁 (2003, 2012), the four variants mentioned above, [dz], [z], [l], and [ɡ], were genealogically involved in a series of sound changes and were derived from one another. As shown in Fig. 1, [dz] was the original realization for /dz/ but has undergone some dialect-dependent simplifications due to its articulatory difficulty. The *Chôan* dialect chose to simplify the affricate into a stop [d], which was later transformed into a lateral [l], following the general weakening trend of /d/ in the language (Ang 洪惟仁 2012; Yang 楊秀芳 1982). On the other hand, the *Chiang* dialect decided to simplify the affricate into a fricative [z], which was later also simplified into [l], a route consistent with Ohala's (1983) predictions. The converging development of the two dialects has made [l] currently the most preferred form. The [ɡ] variant was not related to [l] in its development and was independently derived from [z] in the *Chiang* dialect due to language contact with Hakka. In other words, both /dz/→[z]→[l] and /dz/→[d]→[l] could be regarded as internally motivated sound change routes, while /dz/→[z]→[ɡ] is best deemed as an externally motivated sound change [cf. Hickey (2012)].

Based on evidence from cross-sectional and cross-dialectal data, Ang 洪惟仁 (2003, 2012) concluded that [ɡ] is an unsuccessful attempt by *Chiang*, which had hardly spread to *Chôan* before it waned, while [l] is a successful variant started by *Chôan*, which has been gradually spread to *Chiang*. This explains why [ɡ] only enjoys a minor group of middle-aged *Chiang* users, and few young talkers adopt this variant. It also explains why the transition from [z]/[dz] to [l] is nearly complete only among young *Chôan*, but not *Chiang*, speakers. There does seem to be a correlation between [l] realization and syllable type. Syllables containing an unrounded vowel or glide and those without a nasal ending

Fig. 1 An illustration showing the sound change process of /dz/ proposed by Ang 洪惟仁 (2003, 2012). The *solid arrows* indicate internally motivated sound changes while the *dashed arrow* indicates an externally motivated sound change [cf. Hickey (2012)]. Sounds surrounded by a *single-dotted square* are variants attested in the literature. The *double-dotted square* signals currently the most preferred sound variant. The *block arrow* indicates external language influence. Please see text for explanation

are more resistant to the [l] realization trend. However, Ang 洪惟仁 (2003, 2012) argued that the trend for [l] is irreversible, and it will eventually claim a landslide victory over both [z]/[dz] and [ɡ] for all syllable compositions and become the sole realization for /dz/ in the near future for both *Chiang* and *Chôan*, resulting in /dz/ being completely merged into /l/ in the system.

Ang 洪惟仁 (2012) attributed the motivation for the dominant /dz/→/l/ sound change to three factors. In addition to the above-mentioned universal tendency of avoiding voiced sibilants in languages around the world (Ohala 1983), Ang 洪惟仁 (2012) claimed that both the low functional load of /dz/ in Taiwan Min and the minimal impact on the system imposed by the merger provide additional incentives. In terms of its functional load, /dz/ is relatively restricted in phonotactics in Taiwan Min, as it can only occur before /i j/ and /u w/, but not before other open vowels. Even within the allowed combinations, there exist many gaps. According to the 臺灣閩南語常用詞辭典 *Taiwan Southern Min Common Word Dictionary* (National Languages Committee 國語推行委員會 2011), there are 2927 entries for words containing /si-/, /sj-/, /su-/, or /sw/-initial syllables and 2111 entries for words containing /tsi-/, /tsj-/, /tsu-/, or /tsw/-initial syllables but only 711 entries for words containing /dzi-/, /dzj-/, /dzu-/, or /dzw/-initial syllables, which is roughly a 4:3:1 ratio. In addition, Ang 洪惟仁 (2012) found that all /dz/-initial syllables belong to the category of 讀冊音 *thak-chheh-im*, or "pronunciation of study[f]," which is used mainly in literary contexts and have lower frequency of use in oral communication, the major communication channel for Min[g]. As a result, relatively little confusion would occur in daily usages if /dz/ is to be merged with /l/.

As for the potential impact on the system imposed by the /dz/→/l/ merger, Ang 洪惟仁 (2003, 2012) argued that the system would still be relatively stable despite the change. Taiwan Min includes three types of obstruents in its inventory, stops, affricates, and fricatives, as shown in Table 1. Only stops have all three places of articulation, while affricates and fricatives only appear in the coronal place. Ang 洪惟仁 (2003, 2012) therefore argued that the overall system would still remain relatively stable if /dz/ is to be merged with /l/, as there is no other series in affricates (*Chôan*) or fricatives (*Chiang*) at stake. On the other hand, if any of the stops, say, /b/, is to be merged with another category, say, /p/, then it would subsequently instigate a cascading effect on stops in other places of articulation as well, resulting in major reorganization of

Table 1 The obstruents of Taiwan Min. The dialectal split of *Chôan* and *Chiang* regarding the realization of /dz/ followed Ang 洪惟仁 (2003, 2012)

	Bilabial	Coronal	Dorsal
[−voice, +aspiration]	p^h	t^h	k^h
[−voice, −aspiration]	p	t	k
[+voice, −aspiration]	b		g
[−voice, +aspiration]		ts^h	
[−voice, −aspiration]		ts	
[+voice, −aspiration]		dz (*Chôan*)	
[−voice]		s	
[+voice]		z (*Chiang*)	

phonology, which would then jeopardize the stability of the language and thus encounter considerable resistance from the system. Ang 洪惟仁 (2003, 2012) therefore asserted that the merging of /dz/ with /l/ incurs only minimal cost on Min phonology, and is more likely to be tolerated.

1.1 Specific aims

There are three specific aims in this study. First of all, syllable structure has been observed to be a major influential factor in determining the realization of /dz/ (Ang 洪惟仁 2003). In syllables with a rounded vowel/glide immediately following the onset /dz/, or those ending in a nasal, /dz/ is more likely to be realized as [l] than [z]/[dz]. Ang 洪惟仁 (2003) claimed that this has to do with two intrasyllabic constraints. The first constraint is between voiced fricatives and dorsal sounds. Moving directly from a more marked apical /dz/ to a dorsal /u/ or /w/ requires much effort, and thus a less marked [l] is chosen as the realization target to ease articulation. Analogous situations do not arise in the unrounded environment, as /dz/ is already palatalized via a phonological assimilation rule when it is placed before a palatal /i/ or /j/, and articulation is thus not as effortful. The second constraint is between intrasyllabic fricatives and nasals. Ang 洪惟仁 (2003) used historical evidence to argue that the concurrence of intrasyllabic fricatives and nasals are more marked in Chinese languages in general as it creates inconsistency in stridency, with /dz/ being [+strident] and nasals being [−strident]. By realizing /dz/ as [l], the system thus conforms to the intrasyllabic stridency constraint by making both segments [−strident]. Similar situations do not arise in syllables without nasal endings, as they do not have stridency conflict to begin with. If part of the reason why /dz/ realization is affected by syllabic structures is to ease articulation, then one suspects that the effect of syllable structure might also reflect the gradient nature of syllable complexity. In addition to the major trend observed in Ang 洪惟仁 (2003), one would also expect more [l] and fewer [z]/[dz] realizations in more complex syllables than simpler ones.

Secondly, one would like to investigate the effect of dialect in /dz/ realization among young fluent Min speakers. Ang 洪惟仁 (2003, 2012) predicted that [l] would eventually become the sole realization of /dz/ for the younger generations to come, first for the *Chôan* dialect and later for the *Chiang* dialect. Since Ang 洪惟仁 (2003, 2012) collected his data before 2003, it would be interesting to see whether such predictions are borne out after a time period of over 10 years has passed[h]. If the trajectory of the sound change is in accord with Ang's 洪惟仁 (2003, 2012) projection, then one would expect to see little variation in the realization of /dz/ among different dialects, as speakers would largely realize /dz/ as [l] and only very marginally realize the other two variants, [z]/[dz] and [ɡ], if not forgoing them completely. On the other hand, if the persistent trend of [z]/[dz] found in young speakers of the *Chiang* (Ang 洪惟仁 2003, 2012) and the Mix dialect (Khng 康韶真 2014), and that of [ɡ] found in some young speakers of the Mix dialect (Chen 陳雅玲 2010, 2012) still hold, then one would expect to find a minor yet non-negligible trend of [z]/[dz] among young speakers of the *Chiang* and the Mix dialect, and a robust realization of [ɡ] in at least some young Mix dialect speakers. This would be especially interesting for speakers of the Mix dialect, as Ang 洪惟仁 (2003, 2012) did not provide predictions specific to this group of speakers.

Finally, this study would like to investigate how consistent individual speakers are in realizing /dz/. For sound changes in progress, inter-speaker variability is usually the norm, as individual speakers might be at different stages of a sound change (Janson 1983). However, it is unclear whether individual speakers show consistency when realizing sounds that are undergoing a morphing process. Gósy (2013) observed matched levels of inter- and intra-speaker variability for an ongoing vowel change in Hungarian. However, Yu et al. (2015) found that English-speaking individuals are not likely to differ in their coarticulatory phonetic realizations across time, showing high within-speaker consistency. If the realization of /dz/ in Min is more like the Hungarian vowel change, then one would expect to find large intra-speaker variability, in addition to inter-speaker differences. On the other hand, if Min speakers treat the realization of /dz/ as more like coarticulatory effects found in English, then one would expect that much of the variability in /dz/ realization comes from speakers' various sociolinguistic backgrounds, rather than within-speaker idiosyncratic variations.

Although much research has been done on the realization of Taiwan Min /dz/, it is mainly restricted in two aspects. First of all, except for Ang 洪惟仁 (2003, 2012) and Thoo 涂文欽 (2009), which studied both the *Chiang* and the *Chôan* dialect sampled across various regions in Taiwan, most of the other research focused only on single dialects [*Chôan*: Ang and Chang 洪惟仁, 張素蓉 (2008), Hung 洪慧鈺 (2007), Wang 王薈雯 (2014); Mix: Ang 洪惟仁 (1997), Chen 陳淑娟 (1995), Chen 陳雅玲 (2010, 2012), Khng 康韶真 (2014), Lin 林珠彩 (1995)], and on single locations [Taichung: Ang and Chang 洪惟仁, 張素蓉 (2008); Changhua: Hung 洪慧鈺 (2007); Thoo 涂文欽 (2009); Tainan: Chen 陳淑娟 (1995); Kaohsiung: Ang 洪惟仁 (1997), Chen 陳雅玲 (2010, 2012), Khng 康韶真 (2014), Lin 林珠彩 (1995); Kinmen: Wang 王薈雯 (2014)]. Therefore, generalization could not be easily derived as different studies employed different research methods and elicitation texts, rendering cross-dialectal comparison difficult. This study thus intends to simultaneously compare across the three major dialects of Min, *Chiang*, *Chôan*, and Mix, with regard to the realization of /dz/ in order to provide a more complete picture of the status quo of the sound variation.

Secondly, except for Ang 洪惟仁 (2003, 2012), which used isolated phrases and sentences, in addition to words, and Wang 王薈雯 (2014), which used spontaneous monologue speech, most of the previous studies used syllable or word lists as the main elicitation text (Ang and Chang 洪惟仁, 張素蓉 2008; Chen 陳淑娟 1995; Chen 陳雅玲 2010, 2012; Hung 洪慧鈺 2007; Khng 康韶真 2014; Lin 林珠彩 1995; Thoo 涂文欽 2009). As a consequence, the results might not have fully reflected the true range of variability of /dz/ realization, as syllable and word lists can be a fairly straightforward indicator of the study aim, and speakers would consequently be more conscious of their pronunciation [cf. Fon et al. (2011)]. In other words, speakers might be more inclined to provide the prescribed pronunciations as the preferred variant. In addition, previous studies were mainly conducted via traditional fieldwork methods, without much assistance from instrumental analyses (Ang 洪惟仁 2003, 2012; Ang and Chang 洪惟仁, 張素蓉 2008; Chen 陳淑娟 1995; Chen 陳雅玲 2010, 2012; Hung 洪慧鈺 2007; Khng 康韶真 2014; Lin 林珠彩 1995; Thoo 涂文欽 2009; Wang 王薈雯 2014). Although these studies are of undeniable value, some of the phonetic nuances that are not easily discernible by ear might have been overlooked and potential variances left undocumented due to the transient nature of sounds. This might partially explain why researchers have
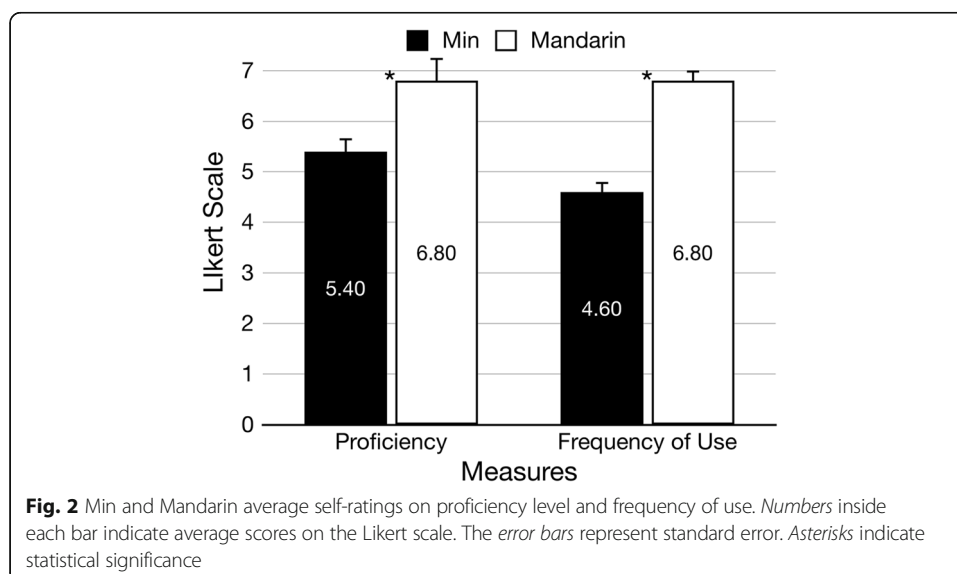
yet to settle the actual realization of some Min sounds that have undergone diachronic or synchronic weakening processes, including /dz/ (Ang 洪惟仁 2003, 2012), /l/ [e.g., Ang 洪惟仁 (2012); Yang 楊秀芳 (1982)], and stops in general (Ratte 2009, 2011). This study thus intends to complement traditional auditory analyses with spectrographic evidence when necessary using longer connected speech to study the realization of /dz/ and examine how syllable complexity and speaker dialect interact with such realizations.

2 Methods

2.1 Participants

Ten fluent Mandarin-Min female bilinguals, aged between 22 and 26 at the time of recording (X = 22.96, SD = 1.34), participated in the experiment. All of them were university students and did not suffer from any speech and language disorders. Except for two, who were Min-first sequential bilinguals that acquired Mandarin only after attending school at the age of five and six, all of the other speakers were simultaneous bilinguals of Min and Mandarin [cf. De Houwer (1995)]. However, since all participants have acquired both languages before age eight, they could be uniformly considered as early bilinguals based on Beardsmore's (1986) criteria.

Participants were asked to self-rate their Min and Mandarin proficiency levels and frequency of use on a 7-point Likert scale, with 1 being the least proficient/frequent and 7 being the most proficient/frequent. As shown in Fig. 2, speakers showed higher proficiency in Mandarin than Min [$t(9)$ = 4.58, p < .001] and used Mandarin much more frequently than Min in their everyday lives [$t(9)$ = 3.71, p < .01]. Min proficiency and Min frequency of use were highly correlated [$r(8)$ = .79, p < .01]. Those who used Min more frequently had higher Min proficiency. Min proficiency and Mandarin frequency of use were also correlated [$r(8)$ = −.67, p < .05]. Speakers who used Mandarin more frequently tended to have lower Min proficiency. There was also a marginal negative correlation between Min and Mandarin frequency of use [$r(8)$ = −.59, p = .07]. Speakers who used Mandarin more frequently were less likely to use Min in their everyday lives.

Fig. 2 Min and Mandarin average self-ratings on proficiency level and frequency of use. *Numbers* inside each bar indicate average scores on the Likert scale. The *error bars* represent standard error. *Asterisks* indicate statistical significance

Despite inter-speaker variability in the self-ratings, it is safe to assume that all speakers had an above-average proficiency level in Min, as none of the speakers self-rated herself as lower than 4, and most speakers self-rated themselves as 5 or above.

The Min dialect group to which each participant belonged was coded according to self-provided demographic and residential information. Both parental dialect and speaker residency were adopted as criteria. Parental dialects were derived from mapping parental hometowns onto Ang's 洪惟仁 (2013) Min dialectal map, and were the main basis for determining the speakers' dialect, as Min is considered to be mainly a language of private domains (Huang 黃宣範 1993), and thus, home should be a major setting for Min usage. In cases where the two parents came from different dialectal areas, speaker residency was used as an operational criterion for dialect determination. Only one speaker was judged according to the second criterion. The 22-year-old female, whose father grew up in the Beidou Township of the Changhua County (a *Chiang* dialect area) and mother grew up in the Lukang Township of the Changhua County (a *Chôan* dialect area), had lived in the Wuri District of Taichung City for 19 years (a *Chiang* dialect area) and was thus coded as belonging to the *Chiang* dialect. A double check on the recording showed that the dialectal coding was congruent with her speech samples. In total, there were four *Chiang* dialect speakers, three *Chôan* dialect speakers, and three Mix dialect speakers. No difference in Min proficiency and frequency of use was found among the three dialects [proficiency: $F(2, 7) = 1.22$, n.s.; frequency: $F(2, 7) = 2.94$, n.s.].

2.2 Materials

A short paragraph embedding six /dz/-initial target syllables was specifically designed for this study. Four of them were followed by /i/ or /j/, and two were followed by /u/ or /w/. In order to avoid potential confound from prosody, all of the syllables chosen were part of a content word and were placed in prosodically prominent positions. Since Min is more commonly a spoken than a written language for most speakers, only highly frequent words were selected to ensure that participants could fluently read aloud the paragraph without much difficulty and awkwardness. As there were few /dz/-initial syllables that were commonly recognized, and one would also like to examine whether the realization of /dz/ is lexicon-specific, three of the syllables were repeated in the paragraph. Two syllables preceding /i/ appeared twice, and one syllable preceding /u/ appeared five times. As a result, there were in total six tokens that preceded /i/ or /j/ and six that preceded /u/ or /w/. To facilitate the reading process of the participants, the paragraph was presented in Chinese characters conforming as closely as possible to the recommended characters promoted by the Ministry of Education in Taiwan (National Languages Committee 國語推行委員會 2011) with minor revisions. Please see the Appendix for the target syllables and the short paragraph.

2.3 Equipment

Recordings were done at a sampling rate of 44,100 Hz using a KORG MR-1000 digital recorder and a Sennheiser HMD 25-1 head-mounted microphone, and were later downsampled to 22,050 Hz using Adobe Audition CS6.

2.4 Procedure

Recording was performed in a sound-treated room. Participants were seated comfortably wearing a head-mounted microphone. They were asked to complete a questionnaire on their language background before the recording started. The experimenter, who was not one of the authors but was a native Mandarin-Min bilingual, checked with each of the speakers to make sure they could fluently and correctly read out the paragraph before the actual recording began. Participants were asked to read the paragraph in a natural manner. Six of the speakers were requested to read the text a second time due to random omission and/or unclear pronunciation of the target syllables[j]. All analyses except for consistency tests (see below) were done only on the last rendition of the reading. The recording session lasted less than 15 min. and speakers were compensated for their participation with a small monetary reward.

2.5 Transcription

The phonetic realizations of all the target syllables were labeled by coupling auditory judgments with spectrographic examination to improve accuracy, especially for phonetic details that are difficult to detect by ear only. Labeling was independently performed by the two authors, both of whom were Mandarin-Min bilinguals as well as trained phoneticians, using the Praat software (Boersma and Weenink 2009). The inter-labeler reliability was fairly high for broad phonetic categories, with a level of agreement of .85 and a Cohen's kappa of .74 ($p < .0001$). As the second author used narrower phonetic categories, the following analyses adopted her labels in order to provide a more complete and detailed picture of /dz/ realization.

3 Results

3.1 Realization of /dz/

Two of the speakers inadvertently omitted the syllable 入 *jip* in 入去 *jip-khì* "to enter." As a consequence, there were 6 (tokens/environment) × 2 (vowel environments) × 10 (speakers) − 2 (omissions) = 118 tokens of /dz/-initial syllables collected, with 58 of them preceding /i/ or /j/ and 60 of them preceding /u/ or /w/. There were in total 15 phonetic variants found for the realization of /dz/, which could be broadly categorized into five types, as shown in Table 2. In general, laterals (hereafter [L]) were the most popular among speakers, accounting for 58% of the data, followed by dental nonsibilants (hereafter [D]) and dental sibilants (hereafter [Z]), which accounted for 20% and 10% of the data, respectively. Together, these three variants constituted 89% of the total

Table 2 Distribution of /dz/ realizations with regard to roundedness. Categories to the left of the vertical line were those that were commonly discussed in the previous literature, while those to the right were not. [Z]: dental sibilants; [G]: velar obstruents; [L]: laterals; [D]: dental nonsibilants; [R]: retroflexes. Please see text for explanation

	[Z]	[G]*	[L]*	[D]*	[R]*	Total
Rounded	4	0	44	5	7	60
Unrounded	8	6	25	19	0	58
Total	12	6	69	24	7	118

*Significant in post hoc tests

data. The retroflex variant (hereafter [R]) was the next most popular, accounting for 6%, followed by the velar obstruents (hereafter [G]), accounting for only 5% of the data. It is interesting to note that the commonly mentioned /dz/ realizations in previous works, [Z], [G], and [L] (Ang 洪惟仁 2003, 2012), only accounted for less than 75% of the realizations. On the other hand, although [D] and [R] categories were little mentioned in the previous literature, they were in fact quite popular among speakers in this study and together accounted for over 25% of the data[k].

The different categories were not evenly distributed with regard to syllabic contexts (Table 2). [G] was exclusively reserved for unrounded environments, while [R] was entirely reserved for rounded ones. Although [Z], [L], and [D] occurred before both rounded and unrounded segments, they did not show the same preferences, either. [L] was more commonly found before rounded segments, while [Z] and [D] were more inclined to occur before unrounded ones. Fisher's exact test was performed [$\chi^2(4) = 28.19$, $p < .0001$], and post hoc analyses using Beasley and Schumacker's (1995) method generally confirmed this observation, showing that [L] and [R] were more likely to occur before rounded segments, while [G] and [D] were more likely to occur before unrounded ones ($p < .001$ for [L] and [D]; $p < .01$ for [R]; $p < .05$ for [G])[l]. As a substantial portion of the tokens were realized with unorthodox pronunciations, and there seemed to be a surprising roundedness effect with regard to the choice of the variants besides the well-documented [L] and [G] [cf. Ang 洪惟仁 (2003, 2012)], one would like to take a closer look at the actual realizations of each broad sound category in the following section in order to provide a more detailed picture with regard to the realization of /dz/.
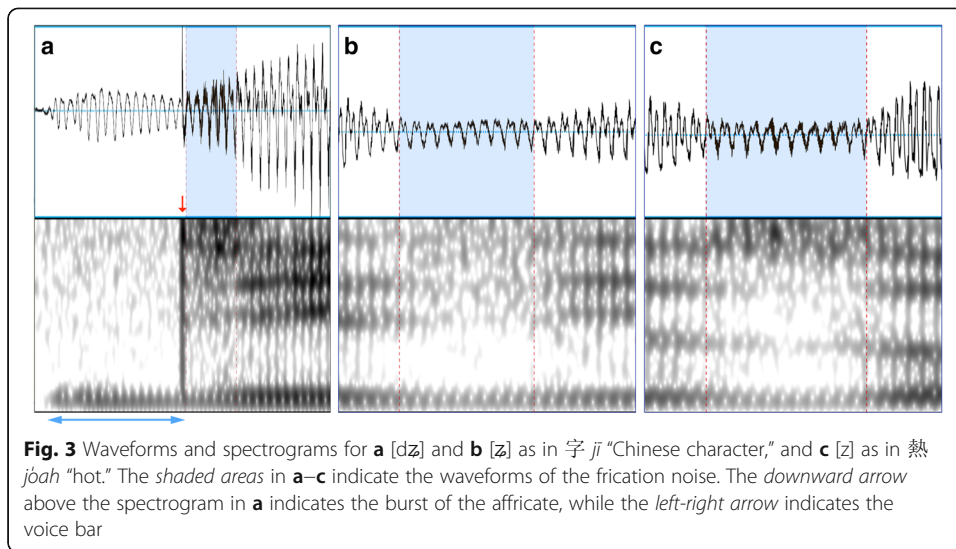
There were 12 tokens that belonged to the [Z] group, including one affricate and two fricatives (Table 3). The two manners of articulation were easily differentiated by spectrographic evidence, as affricates were accompanied by a clear spike showing the stop burst, which was lacking in fricatives (Fig. 3). The affricate occurred only before unrounded segments, which rendered the palatalized [dʑ] variant. For fricatives, there were two variants, [z] and [ʑ], and they were in complementary distribution. [z] occurred before rounded segments, and [ʑ] occurred before unrounded ones. Fisher's exact test was performed [$\chi^2(2) = 10.13$, $p < .01$], and post hoc analyses indicated that roundedness significantly affected the distribution of [z] ($p < .001$) and [ʑ] ($p < .05$), while the distribution of [dʑ] did not reach significance ($p = .16$).

There were six tokens in the [G] group, including [ⁿg] and [ɣ]. At times when it is difficult to discern by ear between stops and their spirantized fricatives [cf. Ratte (2009, 2011)], acoustic spectrograms were used to assist in distinguishing the two, as the former contained a clear spike indicating the stop burst, while the latter did not (Fig. 4). It is also interesting that all of the velar stops found were prenasalized, as could be easily seen by the nasal formants preceding the burst (Fig. 4a). Both [ⁿg] and [ɣ] occurred only before unrounded segments (Table 4). [ⁿg] was contributed by only one

Table 3 Distribution of /dz/ realizations in the [Z] group. Please see text for explanation
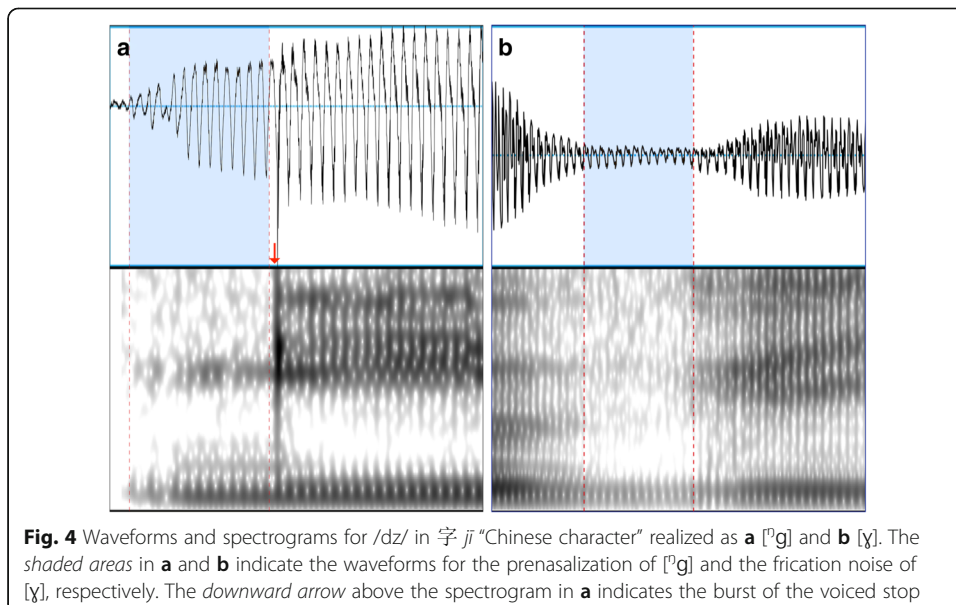
	[z]*	[ʑ]*	[dʑ]	Total
Rounded	4	0	0	4
Unrounded	0	5	3	8
Total	4	5	3	12

*Significant in post hoc tests

Fig. 3 Waveforms and spectrograms for **a** [dẓ] and **b** [ẓ] as in 字 *jī* "Chinese character," and **c** [z] as in 熱 *jòah* "hot." The *shaded areas* in **a**–**c** indicate the waveforms of the frication noise. The *downward arrow* above the spectrogram in **a** indicates the burst of the affricate, while the *left-right arrow* indicates the voice bar

speaker, while [ɣ] came from two speakers. No difference in distribution was found between the two variants [$\chi^2(1) = .67$, n.s.].

The [L] group was the most predominant realization type, which included 69 tokens, accounting for over half of the data. There were three variants found, [l], [ⁿl], and [ʎ] (Table 5). [l] was the most dominant realization, and over 97% of the [L] tokens were realized as such (Fig. 5a). There was also a token of the prenasalized [ⁿl], as could be clearly seen by the energy damping of nasal formants in Fig. 5b, and a token of the palatal lateral approximant [ʎ] (Fig. 5c), both of which were contributed by only one speaker. The talker that produced [ⁿl] was also the one that produced [ⁿg] in the [G] category. Fisher's exact test showed that the distribution of the different variants in the [L] group was not affected by roundedness of the following segment [$\chi^2(2) = 2.19$, n.s.].

Fig. 4 Waveforms and spectrograms for /dz/ in 字 *jī* "Chinese character" realized as **a** [ⁿg] and **b** [ɣ]. The *shaded areas* in **a** and **b** indicate the waveforms for the prenasalization of [ⁿg] and the frication noise of [ɣ], respectively. The *downward arrow* above the spectrogram in **a** indicates the burst of the voiced stop

Table 4 Distribution of /dz/ realizations in the [G] group

	[ᵑg]	[ɣ]	Total
Rounded	0	0	0
Unrounded	4	2	6
Total	4	2	6

The [D] group contained 24 tokens and showed high variability in terms of its realizations, including five different variants in total (Table 6). [d] was the most common, accounting for over 60% of the total [D] tokens, followed by [ˡd], accounting for 21%. The remaining variants, [ⁿd], [ð], and [ˡð], had only one or two tokens each. [d], [ⁿd], and [ð] occurred only before an unrounded segment, while [ˡd] and [ˡð] occurred (mostly) before a rounded one. The five variants were clearly differentiated by spectrographic patterns. The three stop variants were accompanied by a stop burst, which was lacking in the two fricative variants (Fig. 6). Prelateralization and prenasalization were differentiated by the formant pattern before the stop burst. Fisher's exact test showed significance regarding the distribution of the contingency table [$\chi^2(4) = 16.12$, $p < .001$]. Post hoc analyses indicated that roundedness significantly affected the distribution of [d], [ˡd], and [ˡð], generally confirming the observed preferences ($p < .01$ for [d]; $p < .001$ for [ˡd]; $p < .05$ for [ˡð]).
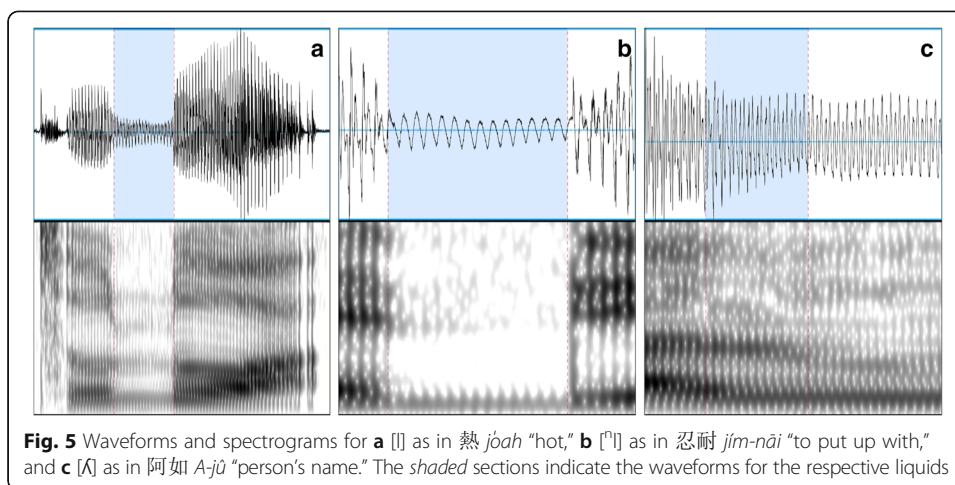
There were seven tokens in total in the [R] group, including two variants, [ʐ] and [ɻ], both of which occurred only before a rounded segment, as shown in Table 7. The two were clearly differentiated by spectrographic patterns. [ʐ] was accompanied by frication noise and a lowered third formant in the following vowel, while [ɻ] was manifested by a lowered third formant during the articulation itself (Fig. 7). [ɻ] was the more common realization, accounting for 86% of the total [R] tokens, while there was only one token of [ʐ]. A chi-square test confirmed this observation [$\chi^2(1) = 3.57$, $p = .059$]. The number of observed cases for [ɻ] was proportionately more than the expected value.

3.2 Realization of /dz/ vs. syllable types

Two syllable types were incorporated in the rounded environment, 如 *jû* and 熱 *joah*, and four syllable types were included in the unrounded environment, 字/二 *jī*, 入 *jip*, 柔 *jiû*, and 忍 *jím*. For rounded syllables, there were four variant categories, [Z], [L], [D], and [R] (Fig. 8a). [L] was the most prominent realization, accounting for over 60% of the data, followed by [R], which accounted for over 10% of the data. [Z] and [D] were relatively minor. Both accounted for no more than 10%. The trend was generally consistent across the two syllables. Fisher's exact test indicated that the overall distribution was not significantly different across the two syllable types, confirming the observation [$\chi^2(3) = 2.25$, n.s.].

Table 5 Distribution of /dz/ realizations in the [L] group

	[l]	[ⁿl]	[ʎ]	Total
Rounded	43	0	1	44
Unrounded	24	1	0	25
Total	67	1	1	69

Fig. 5 Waveforms and spectrograms for **a** [l] as in 熱 *j́oah* "hot," **b** [ⁿl] as in 忍耐 *jím-nāi* "to put up with," and **c** [ʎ] as in 阿如 *A-jù* "person's name." The *shaded* sections indicate the waveforms for the respective liquids

For unrounded syllables, there were also four major realizations, [Z], [G], [L], and [D] (Fig. 8b). Unlike in the rounded environment, the distribution was more varied. [L] was the most dominant in *jím*, accounting for 70% of the data, followed by *jiû* and *jip*. It was the least likely in *jī*, accounting for only 25%. Based on previous studies, it is not surprising to find a predominant proportion of [L] in *jím*, as nasal-ending syllables are observed to be felicitous to such a variant (Ang 洪惟仁 2003, 2012). However, it is interesting that there seemed to be a negative correlation between [L] realization and the phonological duration of the adjacent unrounded segment (i.e., [i] or [j]) for non-nasal-ending syllables. According to Min phonology, the [i] in *jī* has the longest phonological duration since it is in an open syllable, while the [j] in *jiû* is the shortest, as it does not have a vowel status. The [i] in *jip* would be somewhere in-between since it is a syllable bearing an entering tone (Cheng and Cheng Xie 鄭良偉, 鄭謝淑娟 1977; Chung 1997)[m]. It seemed that in non-nasal-ending syllables, [L] was more likely to occur when the following unrounded segment was short. In fact, the [L] ratio for *jiû* was comparable to that of the rounded syllable *joah* (Fig. 8a).
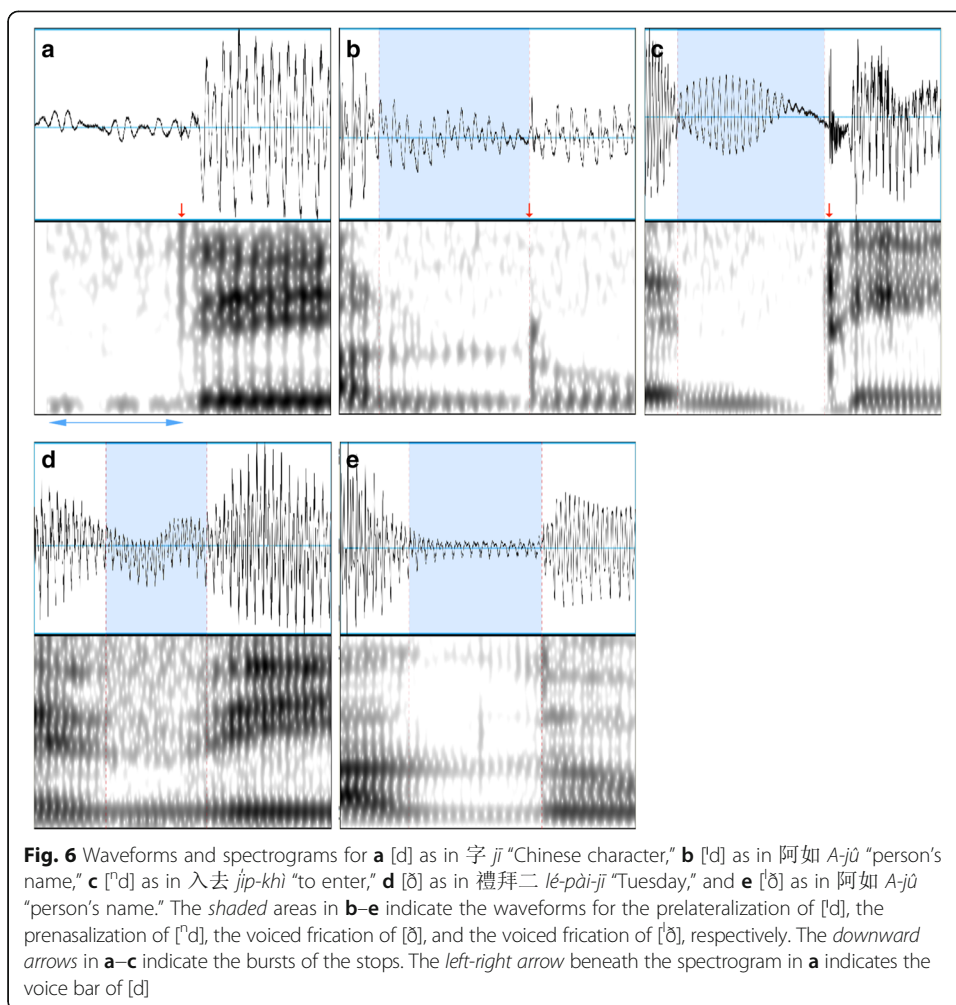
[Z] and [G] were fairly similar in that they preferred open syllables of *jī* and *jiû*, accounting for 25% and 20% for [Z], respectively, and 20% and 10% for [G], respectively. The difference between the two seemed to lie in their avoidance of closed syllables. [Z] was completely absent from the entering tone *jip*, while [G] was not observed in the nasal-ending syllable *jím*. Of the three syllables that showed [Z] realizations, *jī* and *jiû* had both affricate and fricative realizations, while *jím* showed only fricative realizations.

[D] was rather different in terms of its realization pattern. It was the most predominant in *jip*, accounting for 56% of the data, followed by *jī* and *jím*, which had 30% and 20% realization rates, respectively. The syllable *jiû* was the least likely to show such a realization and the rate was as low as 10%. This seemed to imply that [D] prefers

Table 6 Distribution of /dz/ realizations in the [D] group. Please see text for explanation

	[d]*	[ˈd]*	[ⁿd]	[ð]	[ˈð]*	Total
Rounded	0	4	0	0	1	5
Unrounded	15	1	1	2	0	19
Total	15	5	1	2	1	24

*Significant in post hoc tests

Fig. 6 Waveforms and spectrograms for **a** [d] as in 字 *jī* "Chinese character," **b** [ˡd] as in 阿如 *A-jû* "person's name," **c** [ⁿd] as in 入去 *j̍ip-khì* "to enter," **d** [ð] as in 禮拜二 *lé-pài-jī* "Tuesday," and **e** [ˡð] as in 阿如 *A-jû* "person's name." The *shaded* areas in **b**–**e** indicate the waveforms for the prelateralization of [ˡd], the prenasalization of [ⁿd], the voiced frication of [ð], and the voiced frication of [ˡð], respectively. The *downward arrows* in **a**–**c** indicate the bursts of the stops. The *left-right arrow* beneath the spectrogram in **a** indicates the voice bar of [d]

syllables containing the vowel [i], especially those with an entering tone, but tends to avoid syllables containing the glide [j].

Fisher's exact test generally confirmed the above observations. The overall distribution was significant [$\chi^2(9) = 15.67$, $p < .05$]. Post hoc analyses showed that [L] was the most likely to occur in *jím* ($p = .059$) and the least likely to occur in *jī* ($p < .05$). [Z] and [G] were the most common in *jī* ([Z]: $p = .07$; [G]: $p = .08$), and [Z] was the least likely in *jip* ($p < .05$). Finally, [D] was the most likely in *jip* ($p < .05$) and the least likely in *jiû* ($p = .09$).
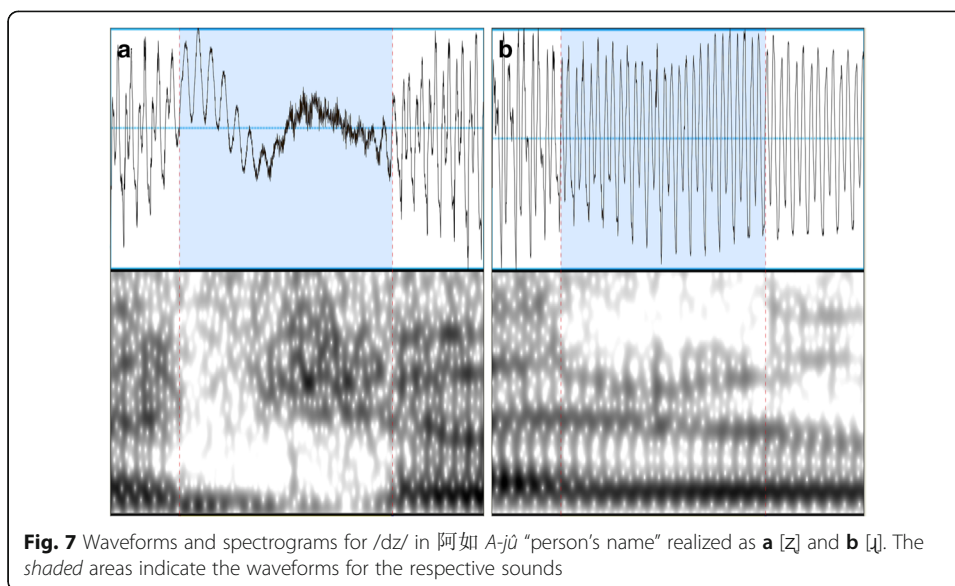
3.3 Realization of /dz/ vs. Min dialects

The effect of dialect was revealed differently in different environments (Fig. 9). For the rounded environment, all three dialects adopted the same set of inventories, [Z], [L],

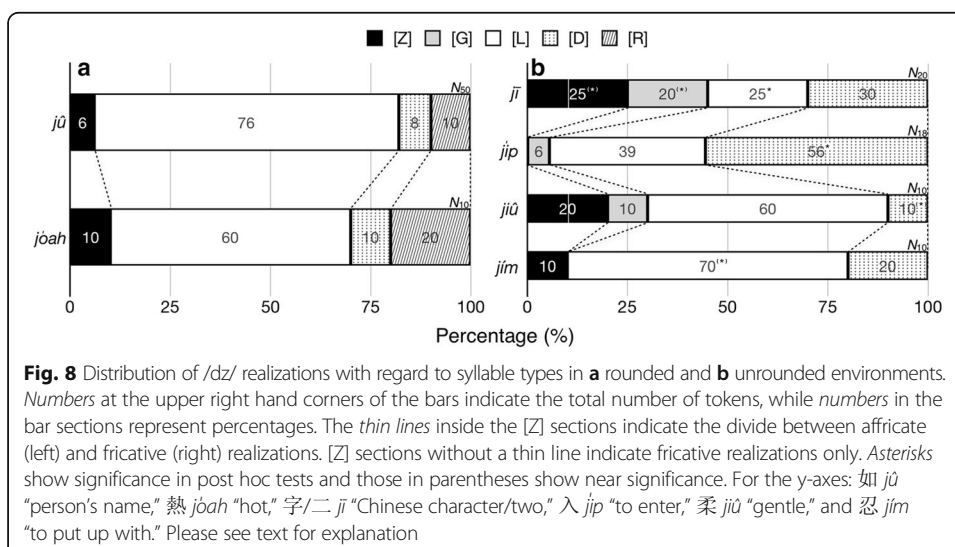**Table 7** Distribution of /dz/ realizations in the [R] group

	[ʐ]	[ɻ]	Total
Rounded	1	6	7
Unrounded	0	0	0
Total	1	6	7

Fig. 7 Waveforms and spectrograms for /dz/ in 阿如 *A-jû* "person's name" realized as **a** [ʐ] and **b** [ɹ]. The *shaded* areas indicate the waveforms for the respective sounds
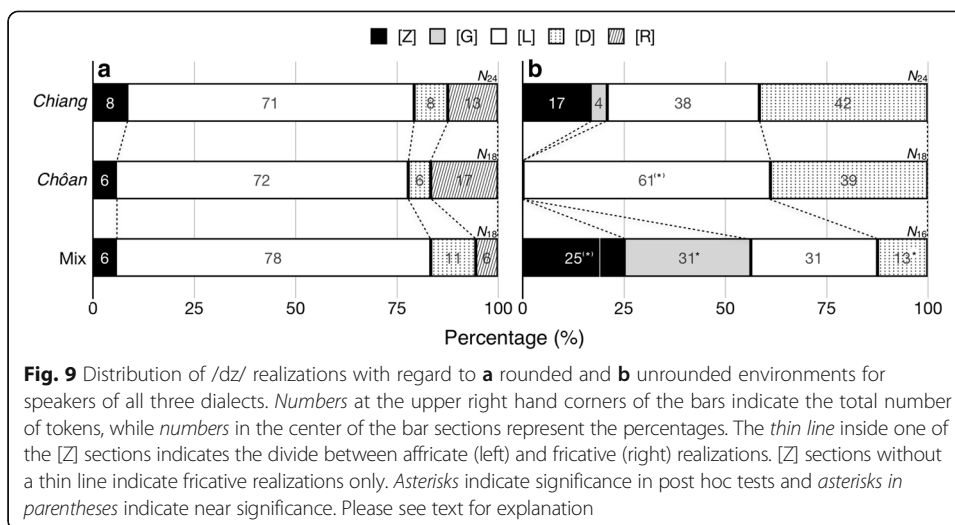
[D], and [R], and also with similar proportions. [L] was the most prominent category in all three dialects, and over 70% of the tokens were realized as such. For *Chiang* and *Chôan*, [R] was the next most common category, accounting for 13% and 17%, respectively, while [Z] and [D] were rather marginal. Less than 10% were realized as such. For the Mix dialect, [D] was the next most common category, accounting for 11%, while [Z] and [R] were rather marginal. Only one token of each was found. However, Fisher's exact test showed that there was no significant dialectal effect with regards to the distribution of the four categories [$\chi^2(6) = 1.64$, n.s.]. All three dialects showed similar tendencies.

In the unrounded environment, both the *Chiang* and the Mix dialect had four variant categories, [Z], [G], [L], and [D]. For *Chiang*, [L] and [D] were the two major realizations, accounting for 38% and 42%, respectively, while [Z] was minor yet robust, accounting for 17% of the total data. [G] was marginal and only one token was
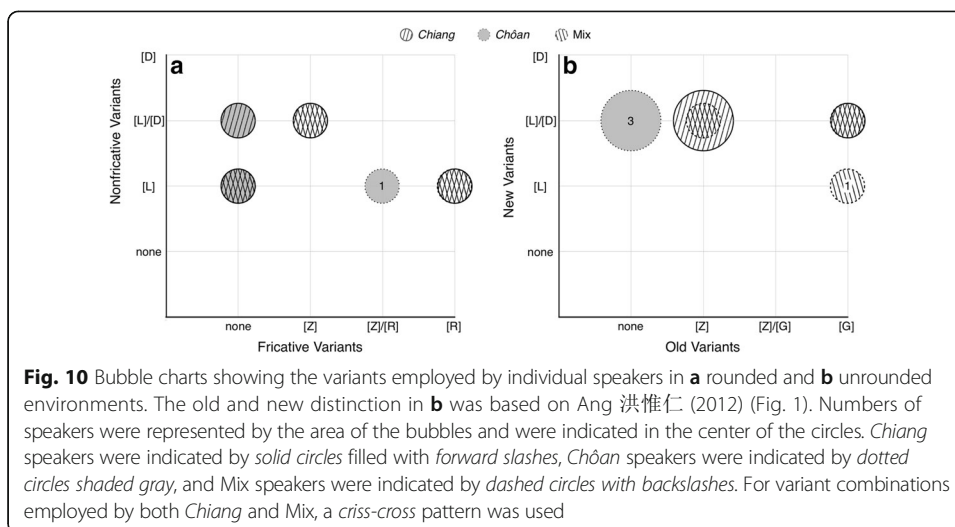

Fig. 8 Distribution of /dz/ realizations with regard to syllable types in **a** rounded and **b** unrounded environments. *Numbers* at the upper right hand corners of the bars indicate the total number of tokens, while *numbers* in the bar sections represent percentages. The *thin lines* inside the [Z] sections indicate the divide between affricate (left) and fricative (right) realizations. [Z] sections without a thin line indicate fricative realizations only. *Asterisks* show significance in post hoc tests and those in parentheses show near significance. For the y-axes: 如 *jû* "person's name," 熱 *jȯah* "hot," 字/二 *jī* "Chinese character/two," 入 *jip* "to enter," 柔 *jiû* "gentle," and 忍 *jím* "to put up with." Please see text for explanation

Fig. 9 Distribution of /dz/ realizations with regard to **a** rounded and **b** unrounded environments for speakers of all three dialects. *Numbers* at the upper right hand corners of the bars indicate the total number of tokens, while *numbers* in the center of the bar sections represent the percentages. The *thin line* inside one of the [Z] sections indicates the divide between affricate (left) and fricative (right) realizations. [Z] sections without a thin line indicate fricative realizations only. *Asterisks* indicate significance in post hoc tests and *asterisks in parentheses* indicate near significance. Please see text for explanation

observed. For the Mix dialect, [Z], [G], and [L] could all be considered as major realizations of /dz/, accounting for 25%, 31%, and 31%, respectively, while [D] remained relatively minor, accounting for only 13% of the data. It is interesting to note that the majority of [Z] were realized as affricates in this dialect, in contrast with *Chiang*, in which only fricative realization was adopted. Relative to these two dialects, *Chôan* simplified its /dz/ realization to only two categories, [L] and [D], with [L] being the major realization, accounting for 61% of the data. Fisher's exact test was performed to confirm the above observations. Results indicated that there was indeed a dialect effect on the realization rates [$\chi^2(6) = 20.82$, $p < .01$]. Post hoc tests showed that *Chôan* had proportionately more [L] ($p = .06$) but fewer [Z] ($p < .05$) and [G] ($p = .08$) than *Chiang* and Mix, while Mix had proportionally more [Z] ($p = .13$) and [G] ($p < .01$) but fewer [D] than *Chiang* and *Chôan* ($p < .05$).

It is interesting to note that no dialect used the same set of variants in its inventory across the two syllable types. During the switch from the unrounded to the rounded environment, both *Chiang* and Mix showed a major increase of [L] realization rates (33% for *Chiang* and 47% for Mix). The increase of [L] was only a modest 11% for *Chôan*, as it is the only dialect that adopted [L] as the predominant category for both rounded and unrounded environments. In addition, both *Chiang* and Mix had a respective 8% and 19% decrease of [Z] and both *Chiang* and *Chôan* showed a 33% decrease of [D]. The effect of the switch on [D] was minimal for the Mix dialect, possibly because the realization rates were already fairly low for this sound category.

Speakers of the same dialect were in general more homogeneous in the unrounded than the rounded environment, as shown in Fig. 10. In the unrounded environment, [L] and [D] were the common variants shared by all *Chiang* speakers. Three of the speakers used an additional [Z] category, while one used an additional [G]. *Chôan* speakers were even more homogeneous, as they all used only [L] and [D], and no other variants were adopted. The Mix speakers were more varied. In addition to [L], which was adopted by all speakers, two speakers adopted [D], two adopted [G], and one adopted [Z], showing more inter-talker variations. Cross-dialectally, [L] and [D] seemed to be the common denominator for almost all speakers except for one, and no speaker simultaneously adopted both [Z] and [G].
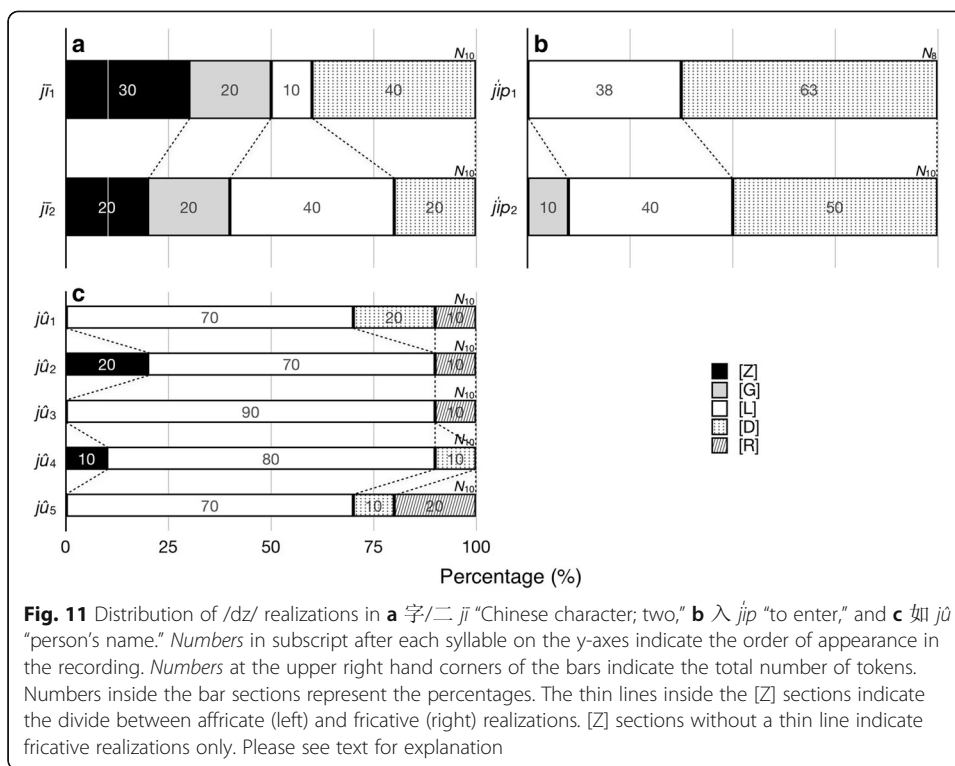
Fig. 10 Bubble charts showing the variants employed by individual speakers in **a** rounded and **b** unrounded environments. The old and new distinction in **b** was based on Ang 洪惟仁 (2012) (Fig. 1). Numbers of speakers were represented by the area of the bubbles and were indicated in the center of the circles. *Chiang* speakers were indicated by *solid circles* filled with *forward slashes*, *Chôan* speakers were indicated by *dotted circles shaded gray*, and Mix speakers were indicated by *dashed circles with backslashes*. For variant combinations employed by both *Chiang* and Mix, a *criss-cross* pattern was used

For the rounded environment, no speaker used all four variants of [Z], [L], [D], and [R]. *Chiang* and Mix speakers behaved rather similarly. Both dialects could simultaneously have as many as three variants for this environment, [Z], [L], and [D]. There were also speakers in both dialects using only [L] and [R], or strictly only [L]. *Chiang* also had one speaker who used only [L] and [D]. For *Chôan* speakers, one could also have as many as three variants for this environment, [Z], [R], and [L]. There was also one speaker who used [L] and [D] and another who used only [L]. Across the three dialects, [L] seemed to be the only variant that was used by all speakers in this environment. No other variant had achieved this status.

3.4 Realization of /dz/ vs. intra-talker consistency

Two analyses were executed to examine intra-talker consistency. One was to see whether speakers realized the same syllables in a similar fashion. The other was to look at whether speakers showed consistent realization patterns when they performed the task twice. In both cases, low variability would imply high intra-speaker consistency despite inter-speaker variation. For the first analysis, there were in total three syllables that had multiple tokens, *jī*, *jip*, and *jû*, and could thus be examined for intra-speaker consistency. There were two repetitions of *jī* and *jip* each and five repetitions of *jû*. Figure 11 shows the distribution of /dz/ realizations for the three syllables.
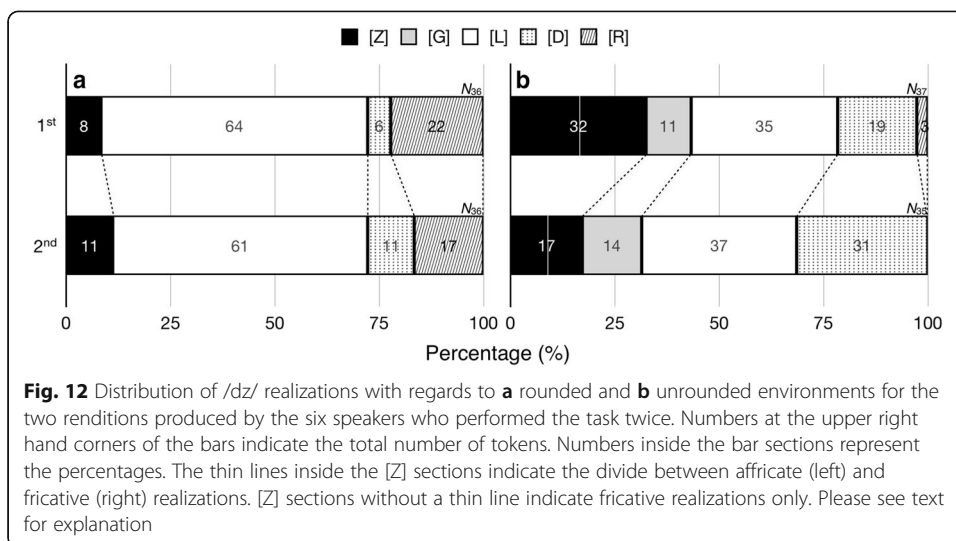
For *jī*, there were four realization categories, [Z], [G], [L], and [D]. No category was overwhelmingly predominant for both renditions, although [D] was more common in the first rendition, while [L] was more common in the second, both accounting for 40% of the data. For *jip*, there were two categories of [L] and [D] for the first rendition, and an additional minor category of [G] for the second. [D] was the more dominant variant for both mentions, accounting for more than 50% of the data, closely followed by [L]. For *jû*, there were four possible categories, [Z], [L], [D], and [R]. However, no rendition simultaneously showed all four. The actual number of realization types was two or three. Despite the variability, all five repeats of *jû* had a predominant [L] realization, accounting for more than 70% of the data, in addition to minor realizations of [Z], [D], and [R]. Generally speaking, there was little variation observed across the repeats of the

Fig. 11 Distribution of /dz/ realizations in **a** 字/二 *jī* "Chinese character; two," **b** 入 *jip* "to enter," and **c** 如 *jû* "person's name." *Numbers* in subscript after each syllable on the y-axes indicate the order of appearance in the recording. *Numbers* at the upper right hand corners of the bars indicate the total number of tokens. Numbers inside the bar sections represent the percentages. The thin lines inside the [Z] sections indicate the divide between affricate (left) and fricative (right) realizations. [Z] sections without a thin line indicate fricative realizations only. Please see text for explanation

same syllables, and this was confirmed by statistical analyses. Three Fisher's exact tests were performed for the three syllables, and none of them revealed significant results [*jī*: $\chi^2(3) = 2.67$, n.s.; *jip*: $\chi^2(2) = .97$, n.s.; *jû*: $\chi^2(12) = 9.39$, n.s.].

Figure 12 shows the /dz/ realizations of the six speakers who recorded the passage twice due to target omission in the first round. For the rounded environment, there were four variants, [Z], [L], [D], and [R]. Regardless of the renditions, [L] was the major realization, accounting for over 60% of the data. [R] was the next largest category, accounting for about 20% of the data. [Z] and [D] were rather marginal, accounting for at most 11%. For the unrounded environment, there were four major realization types, [Z], [G], [L], and [D]. [Z] and [L] were more frequent for the first rendition, both accounting for over 30% of the data, while [L] and [D] were more frequent in the second rendition, both accounting for also more than 30% of the data. It is interesting to note that for both renditions, half of the [Z] variants were realized as affricates and half were realized as fricatives. There was also one instance of [R] found in the first rendition. However, it occurred in a speech error and therefore was likely not a regular realization type for the condition[n]. In general, there was not much difference between the first and the second rendition. Fisher's exact tests confirmed such observation [rounded: $\chi^2(3) = 1.18$, n.s.; unrounded: $\chi^2(4) = 3.88$, n.s.].

Figure 13 shows the variant combinations used by the six individual speakers. It is clear that most speakers adopted similar if not identical variant combinations for both renditions. For the rounded environment, two speakers showed the same combinations across the two trials ([L] and [L]/[D]), while four speakers adopted combinations that only differed by one variant category ([R]/[L] vs. [Z]/[R]/[L]; [R]/[L]/[D] vs. [Z]/[L]/[D];

Fig. 12 Distribution of /dz/ realizations with regards to **a** rounded and **b** unrounded environments for the two renditions produced by the six speakers who performed the task twice. Numbers at the upper right hand corners of the bars indicate the total number of tokens. Numbers inside the bar sections represent the percentages. The thin lines inside the [Z] sections indicate the divide between affricate (left) and fricative (right) realizations. [Z] sections without a thin line indicate fricative realizations only. Please see text for explanation

[Z]/[L] vs. [Z]/[L]/[D]). No speaker used combinations that differed by two or more variants.

The situation was fairly similar in the unrounded environment. Four of the six speakers had exactly the same variant combinations for both trials ([G]/[L]/[D], [L]/[D], and [G]/[L]). One speaker had combinations differing by only one variant ([Z]/[R]/[L] vs. [Z]/[L]/[D]). Since this was the speaker who inadvertently used [R] in this environment due to a speech error, the combination could probably be more accurately reanalyzed as [Z]/[L] vs. [Z]/[L]/[D] instead. There was also one speaker who had combinations that differed by two variants ([Z] vs. [Z]/[L]/[D]).

Fig. 13 Bubble charts showing the variant combinations employed by the six individual speakers who performed the task twice in **a** rounded and **b** unrounded environments. Numbers of speakers are represented by the area of the bubbles and are indicated in the center of the circles. Variant combinations indicated by capital letters on both axes are ordered so that adjacent combinations differ by only one variant. The [R] variant in **b** is in parentheses because it appeared in a speech error. The dashed diagonal line indicates identical variant combinations on both trials

The speaker employed [Z] exclusively for the first rendition, but had one token of [L] and [D] each, in addition to [Z], for the second rendition.

4 Discussion

4.1 Realization of /dz/

The number of /dz/ realizations found in this study, 15 phonetic categories of five broad types, was in strong contrast with the previous literature, as mostly only four were mentioned, [dz], [z], [ɡ], and [l], which, using the notation system adopted in this study, could be further reduced to three broad categories, [Z], [G], and [L] (Ang 洪惟仁 1997, 2003, 2012; Ang and Chang 洪惟仁, 張素蓉 2008; Chen 陳淑娟 1995; Chen 陳雅玲 2010, 2012; Hung 洪慧鈺 2007; Khng 康韶真 2014; Thoo 涂文欽 2009; Wang 王薈雯 2014). The [D] that was found to be fairly robust in the current study was not mentioned before, and [R] was observed only sporadically in previous studies (Chen 陳雅玲 2010, 2012).

The appearance of [D] was the less surprising of the two. Based on sound similarity, [D] was most likely related to [l], both phonologically and phonetically. Phonologically, the voiced dental stop [d] was historically weakened into the dental lateral [l] (Zhang 張振興 1989). In fact, some Min researchers still maintained that /l/ is the phonological counterpart of /b/ and /ɡ/ at the dental place, instead of acknowledging the lack of a voiced dental stop in the system altogether [e.g., Ang 洪惟仁 (2012), Chung (1997), Pan (1995), Yang 楊秀芳 (1982)]. Phonetically, both [d] and [l] are voiced and are produced at the same place. The major difference between the two lies in the existence of a complete oral closure, which only occurs in [d] but not [l]. However, both Yang 楊秀芳 (1982) and Ang 洪惟仁 (2012) argued that the phonetic realizations of the two sounds are even more similar than what is prescribed by the respective phonetic symbols. Yang 楊秀芳 (1982) claimed that [l] in Min is realized as phonetically close to [d], and Ang 洪惟仁 (2012) argued that [l] is produced like a very "soft" [d] before non-low vowels, and is intrinsically a [d]. Although it is unclear what Ang 洪惟仁 (2012) exactly meant by "soft," one speculates that it might have referred to the completeness of the oral closure, as the existence of such is the main distinction between the two sounds. In other words, previous studies seemed to have acknowledged the possibility of [D], and likely deemed [l] and [D] as belonging to the same entity, with Yang 楊秀芳 (1982) claiming an [l] realization with a merged but more [d]-like quality and Ang 洪惟仁 (2012) arguing for a complementary distribution of [l] and [d] being two allophones of the same phoneme. In other words, the existence of [D] might not be new, but rather, it could likely be a variant reflecting the close relationship between [D] and [l] due to historical changes.

However, evidence from the current study seemed to suggest a slightly different picture. First of all, the distribution patterns of [L] and [D] were not the same. [L] more likely occurred in rounded environments, while [D] more likely occurred in unrounded environments (Table 2). The preference of [L] was in accord with the findings in the previous studies (Ang 洪惟仁 2003; Chen 陳淑娟 1995; Lin 林珠彩 1995). Since the /dz/→[L] sound change is occurring at a faster pace in rounded environments than unrounded ones, [L] would naturally be more commonly observed in the former. On the other hand, the distribution of [D] was more intriguing. Based on previous research, there could be two possible predictions. The first prediction is based on Yang's 楊秀芳 (1982) argument of [d] and [l] being the same entity. Based on the claim, [D] and [L] are considered to be

in free variation, and one would expect the preference for rounded environments that was found in [L] to also apply to [D]. However, an opposite trend was observed instead, rendering the free variation possibility unlikely. The second prediction is based on Ang's 洪惟仁 (2012) claim of [l] and [d] being in complementary distribution. According to his proposal, one would expect [l] to more likely occur before low vowels and [d] before non-low vowels. Although such a tendency could not be directly verified in this study, as only high vowels were included, the predominance of [L] realizations in this study rendered this explanation unlikely. In addition, Ang's 洪惟仁 (2012) claim could not explain the differential preference of [D] and [L] regarding roundedness.

Secondly, of the 93 tokens that were identified as either [L] or [D] in this study, there were few "intermediate" tokens found, which should not have been the case if [L] was in fact "phonetically close to [d]" (Yang 楊秀芳 1982) or realized as "a very soft [d]" (Ang 洪惟仁 2012). Most syllables were realized as either a clear [L] (Table 5 and Fig. 5) or a clear [D] (Table 6 and Fig. 6), but not segments of a mixed sound quality. The only segments that could possibly be considered as having a mixed sound quality might be the prelateralized [ˡd]. However, there were only five tokens, accounting for 5% of the total [D] and [L] instances combined, and it was in the opposite direction of what Yang 楊秀芳 (1982) and Ang 洪惟仁 (2012) proposed, as a prelateralized [ˡd] implies a [D] becoming more similar to an [l], instead of an [l] becoming more like a [D], the latter of which is more consistent with the original claim. For [D] to be deemed as equivalent to the [d]-like [l] sound stated in the previous studies, one would need to find instances of prestopped [L] (i.e., [ᵈl]) or at least a semi-closure in the oral cavity of [l] in order to substantiate the claim. However, the [L] category was overwhelmingly realized as typical [l]'s, and crucially, no prestopped [L] was found in the current study (Table 5).

Finally, syllable structure and speaker dialect seemed to play a role in determining whether /dz/ was realized as [L] or [D]. Of the ten speakers recruited in this study, there were nine that had realizations of both [L] and [D], and one that used only [L]. No speaker used only [D]. For those speakers that included both sound categories in their /dz/ realization inventory, whether a token was realized as an [L] or [D] (or other) was largely dependent on both linguistic and social factors. This suggested that the relationship between the two sound categories was not likely to be a straightforward free variation or complementary distribution, as was suggested in the previous studies (Ang 洪惟仁 2012; Yang 楊秀芳 1982), but was instead a reflection of syllabic complexity, speaker background, and/or perhaps even speaker idiosyncrasy.

Even though the [D] category observed in this study is likely a novel form (re-)developed by young speakers of Min, its origin may very well be related to the closeness between [d] and [l] in historical derivations. In other words, speakers might have keenly observed the "void" in the system for a voiced dental stop, and thus (re-)created the [D] realization category to fill the gap. This could be evidenced by the fact that there were some tokens realized as voiced fricatives [ð] and [ˡð] (Table 6), which nicely paralleled the same trend of /b/ being realized as [ʋ] in previous research (Ratte 2009, 2011) and that of [G] being realized as [ɣ] found in this study (Table 4 and Fig. 4). In other words, [D] could possibly be a sound category that was resurrected by young speakers in this study to both fill the missing

gap at the dental position and also create a sound with a constriction more comparable to the original /dz/.

The occurrence of [R] was more surprising. Although Chen 陳雅玲 (2010, 2012) mentioned that a couple of her female speakers showed slight retroflexion in their pronunciations of /dz/ in the rounded environment, it has never been reported as a robust trend. However, approximately 6% of the total tokens collected were realized as such in this study, which was comparable to the percentage of the older [G] variant. If one considers only the rounded environment, in which [R] exclusively occurred, then the percentage even doubled to a non-negligible 12% (Table 2). This indicates that [R] is likely becoming more popular among young speakers, which is intriguing because if decisions for realizing Min /dz/ have been based strictly on phonetic similarity, [R] would not have been a likely top choice. Therefore, some kind of negative transfer from Mandarin must have been at work to have made it happen, which was also suggested by Chen 陳雅玲 (2010, 2012). There were two pieces of evidence to substantiate this claim. First of all, [R] only occurred before rounded /u w/ but was completely absent before /i j/. This follows perfectly from the phonotactic constraints of Mandarin /ʐ/, which is likewise not allowed before high front vowels and palatal glides. Secondly, the two realization variants of the [R] category, [z] and [ɹ], were also found to be the two most common realizations for Mandarin /ʐ/ in Mandarin-Min bilinguals (Chuang et al. 2015).

One possibility why speakers have potentially experienced a negative transfer from Mandarin /ʐ/ in the realization of their Min /dz/ might be due to the similarity in the phonological status and phonetic realization of the two sounds. Both sounds are the only voiced sibilant in the respective systems, and shared a common realization variant [z] (Ang 洪惟仁 2003, 2012; Chan 1984; Chuang et al. 2015). As a consequence, speakers might have made an abstract connection between the two, much like the situation with Min [G] and Hakka /ŋ/ [cf. Ang 洪惟仁 (2012)]. This abstract phonological connection might have then made other phonetic realizations of Mandarin /ʐ/ (i.e., [z] and [ɹ]) become available to Min /dz/ realization. It is of course also possible that the [R] category was a direct negative transfer from Mandarin without the abstract connection, as both of the syllables used in the rounded environment in this study have /ʐ/-initial Mandarin cognates [cf. Ratte (2011)]. Etymologically, 如 *jû* "person's name" is related to Mandarin /ʐu/, while 熱 *joah* "hot" is related to Mandarin /ʐɤə/. It is thus likely that speakers were influenced by the onset in the Mandarin cognates and thus adopted [R] as a novel realization category for /dz/ for these two words. Similar influences from Mandarin were also reported for various Min vowels (Li 李淑鳳 2010; Hsu 許慧如 2016). Min words with vowel variants that are in line with their Mandarin cognates are more preferred than those that are not. In either case, since the majority of [R] was realized as [ɹ], the [R] realization of /dz/ could also to a large extent be qualified as an exemplification of the natural simplification trend proposed by Ohala (1983). Nevertheless, since only two /dz/-initial syllables were included in the rounded environment, more data would have to be collected in order to rule out the possibility of stimulus peculiarity and fully understand the nature of [R] realization in /dz/.

The rich phonetic variability found in this study is rather intriguing and could not be easily discounted as mere speaker idiosyncrasies. The new variants observed might simply be a byproduct of the methodology adopted, as more accurate identification of the actual pronunciation could be achieved when auditory judgment was accompanied by

qualitative spectrographic analyses. It might also be the case that younger Min speakers nowadays indeed adopted more variants in /dz/ realization, possibly due to their lower frequency of use and lower proficiency in Min (Fig. 2). Speakers might be (unconsciously) using more ingenuous ways than their older counterparts to accommodate their declining Min abilities in avoiding the articulatorily challenging [dz]/[z] realizations [cf. Ohala (1983)], as their Min systems might not be as stable. One suspects that reality might be closer to a mixture of the two possibilities combined, with [D] being more likely an overlook due to methodological differences and [R] being more likely a new sound variant due to the predominance of Mandarin usages in the everyday lives of the young people nowadays (Huang 黄宣範 1993). Future studies would be fruitful by incorporating more older speakers as a comparison in order to understand when and how these two sound categories came about, and whether they would remain robust realizations for /dz/ in the years to come.

4.2 Realization of /dz/ vs. syllable types

With regard to the effect of syllable structure, the results of this study seemed to imply an even more complex relationship between /dz/ realization and syllable type than what had previously been suggested. In accord with Ang 洪惟仁 (2003, 2012), /dz/ was indeed more likely to be realized as [L] when it immediately preceded a rounded segment (i.e., *jû* and *joah*) or when it was situated in a syllable with a nasal ending (i.e., *jím*). For syllables such as *jī* and *jip*, the [L] realization rates were largely reduced (Table 2 and Fig. 8). However, a closer inspection of the data showed that *jiû* posed a problem for such a trend. According to Ang 洪惟仁 (2003, 2012), the syllable should be categorized as having a [−round] environment for /dz/ due to the adjacent glide [j], and thus should not have invited much [L] realization. However, the [L] realization rate for *jiû* in this study was in fact fairly comparable to that of rounded environments. In other words, [L] realization seemed not dependent on the roundedness of the immediate following segment, but was more determined by the roundedness of the syllable as a whole. If a syllable contained a rounded segment somewhere, be it contiguous with the onset /dz/ or not, young Min speakers were inclined to realize the onset /dz/ as [L]. One suspects that this might be a reinterpretation and/or expansion of the realization tendency of [L]. As [L] is in the process of becoming the most dominant variant for /dz/, its scope of realization might be enlarged to include other syllabic contexts as well. As a consequence, the motivation for realizing [L] is no longer limited to merely resolving a local articulatory difficulty, as was originally suggested by Ang 洪惟仁 (2003, 2012), but might have instead become a characteristic of a particular syllable type. However, since *jiû* was the only syllable included in this study that showed an exception to Ang's 洪惟仁 (2003, 2012) rules, more data would be needed in order to confirm this.

The realization of [Z] followed a different trend. There was again a syllable type effect within the unrounded environment, in addition to the general roundedness effect (Fig. 8). However, the preferences for [Z] in the unrounded environment were almost the exact opposite of those for [L]. Open syllables of *jī* and *jiû* elicited more [Z] realizations than closed syllables of *jip* and *jím*, and were more likely to realize [Z] as

affricates than fricatives. As Ang 洪惟仁 (2003, 2012) has already observed an avoidance of [Z] in syllables with a final nasal, one suspects that the similar diminishing trend found in *jip* in this study was likely due to analogy. As the general trend was to gradually eradicate [Z] realizations altogether [cf. Ang 洪惟仁 (2003, 2012)], young Min speakers might have thus reduced the scope of [Z] realization to only the simplest syllable structures possible in order to ease articulation.

There was also an effect of syllable type on [G] in the unrounded environment, the only context in which [G] is allowed to occur. Like [Z], [G] preferred open syllables of *jī* and *jiû*, but unlike [Z], it was least likely to be observed in syllables with a final nasal. One suspects that this overall tendency might have to do with the degree of articulatory ease. Since [G] was articulatorily also a challenging sound (Maddieson 2013; Ohala 1983), it is possible that it would prefer syllables with an overall simple makeup, much like the situation with [Z]. In addition, if incongruence in [±dorsal] between the onset and the following segment is one of the major motivations for the /dz/→[L] sound change, as argued by Ang 洪惟仁 (2003, 2012), [G] would be more congruent with the vowel /i/ than [L] because the former two are both [+dorsal], which might explain its relatively robustness in *jī*. For the nasal-ending syllable *jím*, [G] was probably crowded out by the predominance of [L], even though it also contains the vowel /i/. The overall sonorance of the syllable also seemed to be more congruent when [G] was substituted by [L]. In fact, the general distribution trend of [G] was the opposite of [L]. It is likely that contexts suitable for [G] and [L] were complementary. What was deemed as an infelicitous context for [L] was in fact felicitous for [G].

Generally speaking, the /dz/ realization of the young speakers in this study was only partially in accord with Ang's 洪惟仁 (2003, 2012) claims. Talkers indeed showed a structure-dependent realization of [L]. Rounded environments were more felicitous to [L] realization than unrounded environments, and syllables with a final nasal were more felicitous than syllables without. However, the effect of syllabic structure was not only manifested in a binary fashion. For the unrounded environment, which was deemed to be more felicitous to [Z] and [G] realizations, the effect of syllable structure was more of a gradient nature. Syllables of an articulatorily simpler makeup were generally more likely to retain the more marked pronunciations of [Z] ([dz] in particular) and [G] than those that were articulatorily more complex. This implies that the syllabic effect of /dz/ realization for young speakers might not only be phonological, but also phonetic. Whether this phenomenon exists in older speakers is unclear, as Ang 洪惟仁 (2003, 2012) did not look into detailed /dz/ realization of individual syllables. It is possible that speakers of all age groups unanimously demonstrate a gradient syllable-dependent effect in [Z] and [G] realization in the unrounded environment. However, it is also possible that this effect is peculiar to the young speakers only, and is likely a reflection of young speakers' lower Min competence, as negative correlation between proficiency and variability was analogously found in L2 speakers [cf. Nip and Blumenfeld (2015)]. For future research, it would be worth investigating whether Min speakers of various proficiency levels would be differentially affected by syllabic structure in /dz/ realization in unrounded environments, and whether such a pattern could also be observed among older speakers.

4.3 Realization of /dz/ vs. Min dialects

The dialectal effect was found to be structure-dependent. For the rounded environment, the results were generally in accord with Ang's 洪惟仁 (2003, 2012) predictions. Little dialectal variation was observed, and all three dialects unanimously adopted [L] as the predominant category. Although [Z], [D], and [R] were also viable possibilities in this context, they represented a rather minor proportion compared to [L]. Except for [R] in *Chiang* and *Chôan*, which hovered between 10% and 20%, and [D] in Mix, which lingered around 10%, none of the other contexts seemed to be supportive enough for any other non-[L] realizations anymore.

For the unrounded environment, there was much more variability. *Chôan* was qualitatively different from *Chiang* and Mix. There were only two categories observed in *Chôan*, both of which were considered relatively new based on Ang's 洪惟仁 (2012) reconstructions (Fig. 1). [L] was again the predominant realization, but [D] was also fairly robust, accounting for nearly 40% of the tokens. None of the older variants were found in the dialect. On the other hand, *Chiang* and Mix adopted the same set of inventories, incorporating both the newer [L] and [D], and the older [Z] and [G]. The two dialects differed from each other mainly in a quantitative manner. Mix speakers were more inclined to use the older variants than their *Chiang* counterparts, and showed larger proportions of [Z] and [G]. In fact, the two categories were still fairly robust in Mix, together accounting for more than half of tokens. However, only [Z] could be considered as robust in *Chiang*, accounting for 17% of the data, while [G] was merely marginal. In addition, Mix adopted both fricative and affricative realizations for [Z], while *Chiang* only showed fricative realizations. For the newer variants, there was not much difference in the realization of [L], and both dialects showed robust representation of this category. The difference lay in the newest variant [D], which was more popular among *Chiang* than Mix speakers. In other words, except for *Chôan*, Ang's 洪惟仁 (2003, 2012) predictions did not borne out in the unrounded context for the young speakers of this study. The difference between *Chôan* on the one hand, and *Chiang* and Mix on the other regarding the use of [Z] and [L] still persisted in this environment.

Table 8 shows the realization rates of [Z], [G], and [L] in previous research and the current study. Of the three sound categories, [G] was the most consistent with the previous results. As in Ang 洪惟仁 (2003, 2012), young *Chiang* and *Chôan* speakers in this study used very few [G]'s. The frequent usage of [G] found among middle-aged *Chiang* talkers in Ang 洪惟仁 (2003, 2012) was not observed among the young speakers in this study at all. For the Mix dialect, except for Chen 陳淑娟 (1995), which had higher realization rates, the current [G] realization was rather comparable to those in the previous research (Chen 陳雅玲 2010, 2012; Khng 康韶真 2014). This implied that little change has occurred in the past few years for [G]. It has been a fairly stable variant in the Mix dialect and is likely to remain robust for at least a few more years to come. However, young *Chiang* speakers nowadays did not consistently use this variant even though it had been a robust realization once, and it is likely to die out soon in the near future, which is consistent with Ang's 洪惟仁 (2003, 2012) predictions.

More incongruence was found in [Z] and [L]. For [Z], differences were observed mainly in *Chiang* and Mix, while [Z] in *Chôan* remained consistently low in both Ang 洪惟仁 (2003, 2012) and the current study. Compared with Ang 洪惟仁 (2003, 2012), the current *Chiang* speakers had lower [Z] realization rates, but the rates were much

Table 8 Realization rates of [Z], [G], and [L] in previous literature and the current study. Numbers in parentheses indicate combined percentages of [L] and [D]. Chen's 陳雅玲 (2010, 2012) realization rates were indicated by ranges instead of averages because multiple locations were sampled. For studies examining multiple age groups, the age groups that are the most comparable to the speakers in this study were selected. As a consequence, the teenage groups of Ang 洪惟仁 (2003, 2012) and Chen 陳雅玲 (2010, 2012), and the young group of Khng 康韶真 (2014) were used as comparison. Please see text for explanation. +R: rounded environment; −R: unrounded environment

		[Z]		[G]		[L]	
		+R	−R	+R	−R	+R	−R
Chiang	Ang 洪惟仁 (2003, 2012)[a]	21%	31%	—	3%	79%	66%
	This study	8%	17%	—	4%	71% (79%)	38% (79%)
Chôan	Ang 洪惟仁 (2003, 2012)[a]	4%	3%	—	0%	96%	97%
	This study	6%	0%	—	0%	72% (78%)	61% (100%)
Mix	Chen 陳淑娟 (1995)[b]	30%	23%	—	50%	70%	27%
	Chen 陳雅玲 (2010, 2012)	0~26%		0~18%		54~83%	
	Khng 康韶真 (2014)	15%		20%		65%	
	This study	6%	25%	—	31%	78% (89%)	31% (44%)

[a]Ang 洪惟仁 (2003, 2012) included both syllables with and without nasals in his research. However, since this study did not include syllables with nasals in the rounded environment, calculations were based only on syllables without nasals in this context to make the results more comparable
[b]Chen 陳淑娟 (1995) only provided overall data and data for unrounded environments. Data for rounded environments were reconstructed accordingly

lower in the rounded than the unrounded environment. A similar trend was found in the Mix dialect, but only in the rounded environment, in which [Z] was almost completely eradicated. For the unrounded environment, however, the current realization of [Z] was still fairly comparable to that of the previous studies (Chen 陳淑娟 1995; Chen 陳雅玲 2010, 2012; Khng 康韶真 2014). This seemed to imply that the use of [Z] has dwindled in the past decade for both *Chiang* and Mix speakers in a structure-dependent manner. For *Chiang*, the difference in the declining trend is mainly quantitative. Rounded environments showed a higher degree of [Z] elimination than unrounded ones. However, for Mix, there was a qualitative difference. Speakers refrained from using [Z] only in the rounded environment, but persisted in the unrounded environment without any sign of decrease in usage. For *Chôan*, the use of [Z] has already long gone out of fashion, as was argued in Ang 洪惟仁 (2003, 2012), and tokens were found only sporadically. In other words, Ang's 洪惟仁 (2003, 2012) predictions for a complete eradication of [Z] was only fully borne out in *Chôan*. *Chiang* and Mix both showed various degrees of resistance fighting the declining trend.

For [L], the Mix dialect showed results comparable to previous studies (Chen 陳淑娟 1995; Chen 陳雅玲 2010, 2012; Khng 康韶真 2014), while both *Chiang* and *Chôan* showed lower realization rates than those in Ang 洪惟仁 (2003, 2012). For *Chiang*, the major difference lay in the unrounded environment, while for *Chôan*, both the rounded and the unrounded environment showed lower [L] realization. In other words, all three dialects were in direct contradiction with Ang's 洪惟仁 (2003, 2012) predictions, which expected a surge in the use of [L] regardless. One suspects that this might have something to do with the new variants that were adopted by the speakers of the current study, as addition of a new variant would inevitably reduce the proportions of existing categories. It might also be due to the possibility that previous research did not distinguish between [L] and [D], and might have treated the two interchangeably. In fact, if

one combines the [L] and [D] categories together, then the realization rates become more in accord with Ang's 洪惟仁 (2003, 2012) predictions.

Of the three dialects, Mix was surprisingly the most conservative. Among the three dialects, it maintained the highest usage of the older variants [Z] ([dz in particular) and [G], showing the strongest resistance to the general dwindling trend predicted by Ang's 洪惟仁 (2003, 2012). It was also the most reluctant to adopt the new variants of [D] and [R]. On the other hand, *Chôan* was the most liberal. It forwent almost all of the older variants, and readily adopted the newer [L], [D], and [R], showing the smallest effect of roundedness in /dz/ realization. *Chiang* was somewhere in between. It was similar to Mix in persistently retaining the older variant of [Z], albeit with lower frequency of use, but was more like *Chôan* in decisively forsaking [G] and eagerly adopting [D] and [R], showing a balanced taste between tradition and novelty.

The conservative attitude of Mix and, to a lesser extent, *Chiang*, was quite interesting and intriguing, as it did not quite follow Ang's 洪惟仁 (2003, 2012) predictions for young Min speakers' /dz/ realization in specific and was incongruent with other studies regarding young speakers' Min phonology in general (e.g., Li 李淑鳳 2010; Hsu 許慧如 2016). One surmises that this might have to do with a potential sampling bias due to the current distribution of Min speakers. Although all of the speakers recruited in this study showed adequate fluency in Min for carrying out daily conversations in the language and performing the paragraph-reading task smoothly, Min usage and Min fluency of the younger generation at large are in fact on the decline (Directorate-General of Budget, Accounting and Statistics, Executive Yuan 行政院主計總處 2012; The United Daily News Group Poll Center 聯合報系民意調查中心 2002). Therefore, young speakers who are still relatively fluent in Min nowadays are likely to be more conscious about and supportive of Min preservation, and might thus be more conservative in Min usage than their average peers who are less fluent in the language, and speakers of the same age ten or even 20 years ago. This conservative attitude might potentially be the reason why the *Chiang* and Mix speakers had higher-than-expected [Z] and/or [G] realizations.

Regardless of the cause, one surmises that both Mix and *Chôan* would show relative stability at the two respective ends of the spectrum for some time in the near future, while *Chiang* would be more at a transitional stage, possibly shifting towards the *Chôan* end of the spectrum. However, as this study only included a few speakers in each dialect group, and did not incorporate speakers from other age groups, more studies would be needed in order to confirm this.

4.4 Realization of /dz/ vs. speaker consistency

Although speakers could theoretically choose freely from the five broad categories of [Z], [G], [L], [D], and [R] in realizing /dz/, their choices did not seem to be mere random results of idiosyncratic whims, but were constrained in a number of ways. Speakers at most adopted one older variant of [Z] or [G], but used from one to all three of the newer variants of [L], [D], and [R]. If only one variant was employed, then [L] was the obligatory choice. As a result, a single speaker could exclusively use [L] to realize all his /dz/'s throughout, or he could choose to use as many as four variant categories for /dz/ realization.

Speakers were also fairly consistent in their variant choices. There was overall high consistency among different renditions of the same syllables regarding variant choice and distribution (Fig. 11). For speakers that provided repeats of the paragraph, their variant distribution patterns between the two trials were also similar if not identical, both as a group (Fig. 12) and at the individual level (Fig. 13). In other words, even though speakers demonstrated large individual variations due to their different dialectal associations, which also intricately interacted with syllable structure in addition to idiosyncratic choices, intra-speaker variability was in general fairly low. Talkers were rather consistent in their realizations of /dz/ with regard to a particular syllable, and this consistency was maintained when talkers were asked to repeat the task a second time. This implies that /dz/ realization in Min was more in line with the variability of coarticulatory effects found in English (Yu et al. 2015). Even though speakers seemed to show great heterogeneity at the macro-level, with each individual demonstrating his own preferences for /dz/ realization, at the micro-level, each speaker showed high consistency with regard to his decision on /dz/ realization for a particular syllable. Whether this is a trait common to all Min sound variations or whether this is only a peculiar characteristic specific to this variation type awaits further studies.

5 Conclusions

This study examined how /dz/ was realized among fluent young Min speakers in long connected speech using spectrographic evidence as support. Results of this study not only confirmed some of the effects reported in the literature but also added depth to one's understanding of /dz/ realization in specific and sound variation in general. Talkers showed multifarious ways of realizing the articulatorily difficult /dz/ in a nonrandom fashion. The actual realizations that talkers adopted were a complex decision based on a combined concern for syllabic complexity, dialect conservativeness, and idiosyncratic preferences. It is likely that the collective effect of language-internal and language-external factors on the varied realization of /dz/ is not a unique case in Min, but for languages in general. There is a high possibility that all sound variations more or less demonstrate some level of elaborate patterning due to their intertwined relationships with the language and the talker. Future studies on whether other variations follow a similar path and how such variations affect one's percepts would be both interesting and fruitful.

6 Endnotes

[a]The source of the database is http://web.phonetik.uni-frankfurt.de/upsid.html. Accessed 4 June 2015.

[b]Mielke (2007) indicated a language count of 548 in the P-base database. However, the actual number in fact added up to 553 for some unknown reason. The latter was used as the denominator for calculating ratios. The source of the database is http://aix1.uottawa.ca/~jmielke/pbase/. Accessed 4 June 2015.

[c]漢語拼音 Hanyu Pinyin was adopted for Romanization of Mandarin throughout this article.

[d]The 漳 *Chiang* and 泉 *Chôan* dialects are more commonly referred to as *Zhang* and *Quan*, respectively, using Mandarin transliteration. However, this study used the Taiwan Min terms to refer to these two dialects in order to show respect to the speakers of the language. Church Romanization was adopted for Min throughout the article.

[e]Min speakers mainly reside in the west half of Taiwan. Of the two dialects, *Chiang* is more dominant than *Chôan* in Taiwan. However, most speakers show various degrees of *Chiang-Chôan* mixture due to frequent travels within the island (Ang 洪惟仁 1985).

[f]Like many Chinese languages, a great number of Taiwan Min words have two context-dependent pronunciations (Yang 楊秀芳 1982). 讀冊音 *Thak-chheh-im* is used in literary contexts, while 講話音 *kóng-ōe-im*, or "pronunciation of speech," is used in colloquial contexts.

[g]Except for a few Min scholars and writers, and congregations affiliated with the Presbyterian Church in Taiwan, which uses the Min Bible and Min Hymns in the majority of its services, most Min speakers use Min only in oral communication and are not familiar with the written form of Min. According to the statistics provided by the Presbyterian Church in Taiwan 台灣基督長老教會總會 (2013), its congregations accounted for a little over 1% of the total population in 2013.

[h]Ang 洪惟仁 (2003) and Ang 洪惟仁 (2012) presented the same set of data. According to Ang 洪惟仁 (2003), the data were mainly collected between 1999 and 2002, which was approximately 14–17 years ago.

[i]Although the Ministry of Education in Taiwan has recommended a character-based Min transcription system (National Languages Committee 國語推行委員會 2011), some of the characters were not intuitively pronounceable by young native speakers in a pretest. In order to facilitate smooth elicitation, alternatives based on a make-do system widely used in Min lyrics and creative writing were adopted for these specific characters. These included *gín-á* "child," for which 囝仔 was used instead of 囡仔; *m̄-koh* "but," for which 嘸擱 was used instead of 毋過; *kā* "object marker," for which 甲 was used instead of 共; *siūⁿ-beh* "want to," for which 想嘜 was used instead of 想欲; and *hō͘* "to let," for which 乎 was used instead of 予. Talkers showed little trouble reading the passage after the revision.

[j]Since reading Min out loud is not an activity commonly performed by Min speakers, many of the participants tended to paraphrase what they read instead of reading verbatim for the first pass. As a consequence, they were asked to read again, as many of the omissions/alternations were in fact target words.

[k]Chen 陳雅玲 (2010, 2012) did note that a couple of her middle-aged female speakers showed slight retroflexion in some miscellaneous tokens.

[l]Post hoc analyses for chi-square tests were executed using Beasley and Schumacker's (1995) method throughout the paper.

[m]An entering tone is a syllable ending in [p, t, k, ʔ] (Cheng and Cheng Xie 鄭良偉, 鄭謝淑娟 1977; Chung 1997).

[n]The speaker mispronounced the /i/ vowel in 入 *jip* "to enter" as [u], changing the vowel quality from unrounded to rounded.

7 Appendix

Table 9 The recording material of this study. Target syllables are in underline, and the phrasing reflects the original punctuation

Text	Romanization	Translation
有一個囝仔叫做阿 如,	Ū chi̍t-ê gín-á kiò-chò A-jû,	There was a child named A-Ju.
伊足溫柔,	i chiok un-jiû,	She was very sweet,
嘛擱字寫甲足歹,	m̄-koh jī siá kah chiok bái,	but she has poor handwriting,
所以伊的同學攏甲伊笑。	só·-í i ê tông-o̍h lóng kā i chhiò.	so her classmates made fun of her.
頂禮拜二,	Téng-lé-pài-jī,	Last Tuesday,
天氣足熱,	thiⁿ-khì chiok jo̍ah,	it was very hot,
阿如想唛入去教室,	A-jû siūⁿ-beh ji̍p-khì kàu-sek,	and A-Ju wanted to go into the classroom,
嘛擱伊同學甲伊欺負,	m̄-koh i tông-o̍h kā i khi-hū,	but her classmate bullied her
不要乎阿如入去。	bô-ài hō· A-jû ji̍p-khì.	by not letting her in.
阿如心內足艱苦,	A-jû sim-lāi chiok kan-khó·,	A-Ju was in anguish,
所以轉去甲媽媽講。	só·-í tńg-khì kā ma-ma kóng.	so she went home and told her mom.
媽媽就甲伊講:	Ma-ma tio̍h kā i kóng,	Mom then told her,
「阿如,	"A-jû,	"A-Ju,
你就要忍耐!」	lí tio̍h-ài jím-nāi!"	you just have to put up with it!"

Author details
[1]Graduate Institute of Linguistics, National Taiwan University, 1 Roosevelt Rd. Sec. 4, 317 Lexue Bldg., Taipei 10617, Taiwan, ROC. [2]Taiwan International Graduate Program on Computational Linguistics and Chinese Language Processing, Academia Sinica, Taipei, Taiwan, ROC. [3]Graduate Institute of Brain and Mind Sciences, National Taiwan University, Taipei, Taiwan, ROC.

References
Ang, Uijin 洪惟仁. 1985. A study on Taiwan Holo tone 臺灣河佬話聲調研究. Taipei: Zili Wanbao.
Ang, Uijin 洪惟仁. 1997. Southern Min dialects in Kaohsiung county 高雄縣閩南語方言. Kaohsiung: Kaohsiung County Government.
Ang, Uijin 洪惟仁. 2003. The motivation and direction of sound change: On the competition of Minnan dialects Chang-chou and Chuan-chou and the emergence of general Taiwanese 音變的動機與方向:漳泉競爭與台灣普通腔的形成. Ph.D. dissertation. Hsinchu, Taiwan: National Tsing Hua University.
Ang, Uijin 洪惟仁. 2005. Looking at the changes of Taiwan Southern Min from dialect maps graphed in two different eras 從兩個時期製作的方言地圖看台灣閩南語的變化. Fuzhou, China: Paper presented at the 9th Min Dialect International Symposium 第九屆閩方言國際學術研討會.
Ang, Uijin 洪惟仁. 2012. The drift of change of the initial /j-/ of Southern Min 閩南語入字頭(日母)的音變潮流. Journal of Taiwanese Languages and Literature 臺灣語文研究 7:1–32. Available at http://www.uijin.idv.tw/tlls/jtll/abstract/7.2-1.pdf. Accessed 4 June 2015.

Ang, Uijin 洪惟仁. 2013. The distribution and regionalization of varieties in Taiwan 台灣的語種分布與分區. Language and Linguistics 語言暨語言學 14(2):313–367. Available at http://www.ling.sinica.edu.tw/files/publication/j2013_2_05_2724.pdf. Accessed 4 June 2015.

Ang, Uijin, and Su-Rong Chang 洪惟仁, 張素蓉. 2008. The gradient distribution of Quanzhou dialect in the Haixian area of the Taichung county: A socio-geodialectological study 台中縣海線地區泉州腔的漸層分布—— 一個社會地理方言學的研究. In Proceeding of Sociolinguistics and Functional Grammar 社會語言學與功能語法論文集, ed. Hsu S. Wang and Fu-mei Hsu王旭, 徐富美, 13–43. Taipei: Crane. Available at http://www.ntcu.edu.tw/ogawa/history/1st/1-4.pdf. Accessed 4 June 2015.

Balise, Raymond R, and Randy L Diehl. 1994. Some distributional facts about fricatives and a perceptual explanation. Phonetica: International Journal of Speech Science 51(1–3):99–110.

Beardsmore, Hugo B. 1986. Bilingualism: Basic principles, 2nd ed. Clevedon, UK: Multilingual Matters.

Beasley, TMark, and Randall E Schumacker. 1995. Multiple regression approach to analyzing contingency tables: Post hoc and planned comparison procedures. The Journal of Experimental Education 64(1):79–93.

Boersma, Paul, and David Weenink. 2009. Praat: Doing phonetics by computer (Version 5.1) [Computer Program].. Available at http://www.praat.org/. Accessed 28 Aug 2014.

Chan, Hui-chen. 1984. The phonetic development of Mandarin /z/ in Taiwan: A sociolinguistic study. M.A. thesis. Taipei: Fu Jen Catholic University.

Chen, Su-Chuan 陳淑娟. 1995. A study on the sound change from the chu-rhyme to the shi-rhyme in Guanmiao dialect 關廟方言「出歸時」的研究. M.A. thesis. Taipei: National Taiwan University.

Chen, Ya-ling 陳雅玲. 2010. Phonetic variation in Chuanchou-accented Southern Min as spoken in coastal Kaohsiung City 高雄市海岸地帶偏泉腔閩南語的語音變異. M.A. thesis. Hsinchu, Taiwan: National Hsinchu University of Education.

Chen, Ya-ling 陳雅玲. 2012. The variation and change of j-initial syllables in Kaohsiung Southern Min 高雄市閩南語入字頭的變異與變化. Paper presented at the 9th Workshop on the Relationship between the Racial Migration and Distribution of Languages or Dialects 第九屆語言文化分佈與族群遷徙工作坊, Taipei. Available at http://www.uijin.idv.tw/ogawa/history/9th/9-3.pdf. Accessed 4 June 2015.

Cheng, Robert L, and Susie S Cheng Xie 鄭良偉, 鄭謝淑娟. 1977. Phonological structure and romanization of Taiwanese Hokkien 台灣福建話的語音結構及標音法. Taipei: Student Book Co., Ltd.

Chuang, Ya-Wen, Chong-Wei Feng, and Ru-Yi Chen 莊雅雯, 馮鐘緯, 陳如億. 2009. The differences of the variants of j-initial syllables between Helauke and non-Helauke areas 〈入〉字頭「g」變體在鶴佬客地區與非鶴佬客地區之差異. Paper presented at the 3rd Workshop on the Relationship between the Racial Migration and Distribution of Languages or Dialects第三屆台灣的語言方言分佈與族群遷徙工作坊, Kaohsiung, Taiwan. Available at http://www.ntcu.edu.tw/ogawa/history/3rd/3-9.pdf. Accessed 4 May 2015.

Chuang, Yu-Ying, Sheng-Fu Wang, and Janice Fon. 2015. Cross-linguistic interaction between two voiced fricatives in Mandarin-Min simultaneous bilinguals. In Proceedings of the 18th International Congress of Phonetic Sciences, ed. The Scottish Consortium for ICPhS 2015, Paper number 0311. Glasgow, UK: The University of Glasgow. Available at https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0311.pdf. Accessed 4 Apr 2016.

Chung, Raung-fu. 1997. The segmental phonology of Southern Min in Taiwan. Taipei: Crane.

De Houwer, Annick. 1995. Bilingual language acquisition. In The handbook of child language, ed. Paul Fletcher and Brian MacWhinney, 219–250. Oxford: Wiley-Blackwell.

Directorate-General of Budget, Accounting and Statistics, Executive Yuan 行政院主計總處. 2012. The 2010 population and housing census: A general summary report on the statistic results and analyses 99年人口及住宅普查總報告提要分析. Taipei: Directorate-General of Budget, Accounting and Statistics, Executive Yuan. Available at http://www.dgbas.gov.tw/ct.asp?xItem=31969&ctNode=3272&mp=1. Accessed 4 June 2015.

Fon, Janice, Jui-mei Hung, Yi-Hsuan Huang, and Hui-ju Hsu. 2011. Dialectal variations on syllable-final nasal mergers in Taiwan Mandarin. Language and Linguistics 12(2):273–311. Available at http://www.ling.sinica.edu.tw/files/publication/j2011_2_01_2075.pdf. Accessed 4 June 2015.

Gósy, Mária. 2013. Inter-speaker and intra-speaker variability indicating a synchronous speech sound change. In VL1xx: Papers in linguistics presented to László Varga on his 70th birthday, ed. Péter Szigetvári, 313–332. Budapest: Tinta Publishing House. Available at http://seas3.elte.hu/VLlxx/gosy.html. Accessed 4 June 2015.

Gussenhoven, Carlos, and Rolf H. Bremmer Jr. 1983. Voiced fricatives in Dutch: Sources and present-day usage. North-Western European Language Evolution 2:55–71.

Hickey, Raymond. 2012. Internally and externally motivated language change. In The handbook of historical sociolinguistics, ed. Juan Manuel Hernández-Compoy and Juan Camilo Conde-Silvestre, 387–407. Malden, MA: Wiley-Blackwell.

Holes, Clive. 1990. Gulf Arabic. London: Routledge.

Hsu, Hui-ju 許慧如. 2016. The dynamics of Taiwanese–The analysis of the Taiwanese phoneme /o/ 變動中的台語:台語/o/音素三種主要讀音的現狀分析. Soochow Journal of Chinese Studies 東吳中文學報 31:303–328.

Huang, Shuanfan 黃宣範. 1993. Language, society, and ethnic identity: A study on language sociology in Taiwan 語言, 社會與族群意識: 台灣語言社會學的研究. Taipei: Crane.

Hung, Hui-yu 洪慧鈺. 2007. An investigation and analysis on the sound change of j-initial syllables in Fangyuan Township, Changhua County 彰化縣芳苑鄉〈入〉字頭音變的調查與分析. Annual of Graduate School of Chinese Literature Soochow University 有鳳初鳴年刊 3:119–133. Available at http://tao.wordpedia.com/show_pdf.ashx?sess=yxd4xz45wqncqz55kisjxs45&file_name=JO00001287_3_119-133&file_type=q. Accessed 4 June 2015.

Janson, Tore. 1983. Sound change in perception and production. Language 59(1):18–34.

Jesus, Luis MT, and Christine H. Shadle. 2003. Temporal and devoicing analysis of European Portuguese fricatives. In Proceedings of the 15th International Congress of Phonetic Sciences, ed. Maria-Josep Solé, Daniel Recasens, and Joaquín Romero, 779–782. Barcelona: Causal Productions. Available at https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_0779.pdf. Accessed 4 June 2015.

Khng, Siau-Tsin 康韶真. 2014. A phonetic survey and variation analyze of Taiwanese in Kaohsiung 高雄台語語音的調查佮演變分析. M.A. thesis. Taipei: National Taiwan Normal University.

Li, Chung-min 李仲民. 2009. A retrospective and perspective view on the postwar geolinguistics of Taiwan Southern Min 戰後臺灣閩南語地理語言學的回顧與展望. Journal of Taiwanese Languages and Literature 臺灣語文研究 4:107–151.

Li, Siok-Hong 李淑鳳. 2010. Sound change in Taiwanese: A endency caused by its contact with Mandarin 臺、華語接觸所引起的台語語音的變化趨勢. Journal of Taiwanese Vernacular 台語研究 2(1):56–71.

Lin, Ju-Tsai 林珠彩. 1995. A preliminary investigation on the phonetics of some lexical items comparing across three generations of Taiwan Min speakers–Using the Lin family in Xiaogang, Kaohsiung City as an example 台灣閩南語三代間語音詞彙的初步調查與比較——以高雄市小港區林家為例. M.A. thesis. Taipei: National Taiwan Normal University.

Maddieson, Ian. 1984. The UCLA phonological segment inventory database. Available at http://web.phonetik.uni-frankfurt.de/upsid.html. Accessed 4 June 2015.

Maddieson, Ian. 2013. Voicing and gaps in plosive systems. In The world atlas of language structures online, ed. Matthew S. Dryer and Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available at http://wals.info. Accessed 4 June 2015.

Mielke, Jeff. 2007. P-base. Available at http://aix1.uottawa.ca/~jmielke/pbase/. Accessed 4 June 2015.

Miller, George A, and Patricia E Nicely. 1955. An analysis of perceptual confusions among some English consonants. Journal of Acoustical Society of America 27(2):338–352.

National Languages Committee 國語推行委員會. 2011. The Taiwan Southern Min common word dictionary 臺灣閩南語常用詞辭典. Taipei: Ministry of Education, Taiwan. Available at http://twblg.dict.edu.tw/holodict_new/index.html. Accessed 4 June 2015.

Nip, Ignatius SB, and Henrike K Blumenfeld. 2015. Proficiency and linguistic complexity influence speech motor control and performance in Spanish language learners. Journal of Speech, Language, and Hearing Research 58(3): 653–668.

Norman, Jerry. 1988. Chinese. Cambridge: Cambridge University Press.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In The production of speech, ed. Peter F. MacNeilage, 189–216. New York: Springer-Verlag. Available at http://linguistics.berkeley.edu/~ohala/papers/macn83.pdf. Accessed 4 June 2015.

Pan, Ho-hsien. 1995. The phonetic variants of Taiwanese "voiced" stops: An airflow study. In Proceedings of the ROCLING VIII, 191–215. Available at https://aclweb.org/anthology/O/O95/O95-2004.pdf. Accessed 4 June 2015.

Ratte, Alexander T. 2009. A dialectal and phonological analysis of Penghu Taiwanese. B.A. thesis. Williamstown, USA: Williams College. Available at http://sanders.phonologist.org/Papers/ratte-thesis.pdf. Accessed 4 June 2015.

Ratte, Alexander T. 2011. Contact-induced phonological change in Taiwanese. M.A. thesis. Columbus, USA: The Ohio State University. Available at https://etd.ohiolink.edu/!etd.send_file?accession=osu1313497239&disposition=inline. Accessed 4 June 2015.

Schuster-Šewc, Heinz. 1999. Grammar of the Upper Sorbian language: Phonology and morphology. München: Lincom Europa.

Singh, Sadanand, and John W. Black. 1966. Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. Journal of Acoustical Society of America 39(2):372–387.

Smith, Caroline L. 1997. The devoicing of /z/ in American English: Effects of local and prosodic context. Journal of Phonetics 25(4):471–500.

The Presbyterian Church in Taiwan 台灣基督長老教會總會. 2013. The congregation population of the congregation in 2013 2013年信徒總數. Available at http://churchstat.pct.org.tw/datalist.htm. Accessed 4 June 2015.

The United Daily News Group Poll Center 聯合報系民意調查中心. 2002. The legacy and disappearance of mother tongues 母語的傳承與流失, United Daily News, 14. Available at http://www.hakkaonline.com/thread-5123-1-1.html. Accessed 4 June 2015.

Thoo, Bun-khim 涂文欽. 2009. A geo-dialectological study on Southern Min in Changhua County 彰化縣閩南語地理方言學研究. Paper presented at The 10th National Conference on Linguistics 第十屆全國語言學論文研討會, Taoyuan, Taiwan. Available at http://fl.hs.yzu.edu.tw/NCL2009/papers/G21.pdf. Accessed 4 June 2015.

Tinelli, Henri. 1981. Creole phonology. The Hague: Mouton.

Wang, Hui-Wen 王薈雯. 2014. The performances of the Kinmen dialect of young people in Kinmen using 8 young people born in Kinmen in the 80s as examples 金門青年的金門話表現情形:以八位1980年代出世的金門青年做例. M.A. thesis. Taipei: National Taiwan Normal University.

Wang, Marilyn D, and Robert C Bilger. 1973. Consonant confusions in noise: A study of perceptual features. Journal of Acoustical Society of America 54(5):1248–1266.

Yang, Hsiu-Fang 楊秀芳. 1982. A study on the literary-colloquial system of Southern Min 閩南語文白系統的研究. Ph.D. dissertation. Taipei: National Taiwan University.

Yao, Rong-song 姚榮松. 1988. The phonemic system and the nasalized finals of the Southern Min rhyme book: Xuei-in-miao-wu 彙音妙悟的音系及其相關問題. Bulletin of Chinese 國文學報 17:251–258.

Yu, Alan CL, Carissa Abrego-Collier, Jacob Phillips, Betsy Pillion, and Daniel Chen. 2015. Investigating variation in English vowel-to-vowel coarticulation in a longitudinal phonetic corpus. In Proceedings of the 18th International Congress of Phonetic Sciences, ed. The Scottish Consortium for ICPhS 2015, paper number 0519. Glasgow, UK: The University of Glasgow. Available at https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0519.pdf. Accessed 4 Apr 2016.

Zhang, Zhenxing 張振興. 1989. Notes on Southern Min dialects 臺灣閩南方言記略. Taipei: Wen Shi Zhe.

Żygis, Marzena. 2008. On the avoidance of voiced sibilant affricates. ZAS Papers in Linguistics 49:23–45. Available at http://www.zas.gwz-berlin.de/fileadmin/material/ZASPiL_Volltexte/zp49/zaspil49-zygis.pdf. Accessed 4 June 2015.

Żygis, Marzena, Susanne Fuchs, and Laura L König. 2012. Phonetic explanations for the infrequency of voiced sibilant affricates across languages. Laboratory Phonology 3(2):299–336. Available at http://www.haskins.yale.edu/Reprints/HL1715.pdf. Accessed 4 June 2015.