The Southwestern Mandarin /n/-/l/ Merger: Effects on Production in Standard Mandarin and English

Wei Zhang[1] and John M. Levis[2]*

[1]*School of Translation, Qufu Normal University, Qufu, China,* [2]*Department of English, Iowa State University, Ames, IA, United States*

Southwestern Mandarin is one of the most important modern Chinese dialects, with over 270 million speakers. One of its most noticeable phonological features is an inconsistent distinction between the pronunciation of (n) and (l), a feature shared with Cantonese. However, while /n/-/l/ in Cantonese has been studied extensively, especially in its effect upon English pronunciation, the /l/-/n/ distinction has not been widely studied for Southwestern Mandarin speakers. Many speakers of Southwestern Mandarin learn Standard Mandarin as a second language when they begin formal schooling, and English as a third language later. Their lack of /l/-/n/ distinction is largely a marker of regional accent. In English, however, the lack of a distinction risks loss of intelligibility because of the high functional load of /l/-/n/. This study is a phonetic investigation of initial and medial (n) and (l) production in English and Standard Mandarin by speakers of Southwestern Mandarin. Our goal is to identify how Southwestern Mandarin speakers produce (n) and (l) in their additional languages, thus providing evidence for variations within Southwestern Mandarin and identifying likely difficulties for L2 learning. Twenty-five Southwestern Mandarin speakers recorded English words with word initial (n) and (l), medial <ll> or <nn> spellings (e.g., swallow, winner), and word-medial (nl) combinations (e.g., only) and (ln) combinations (e.g., walnut). They also read Standard Mandarin monosyllabic words with initial (l) and (n), and Standard Mandarin disyllabic words with (l) or (n). Of the 25 subjects, 18 showed difficulties producing (n) and (l) consistently where required, while seven (all part of the same regional variety) showed no such difficulty. The results indicate that SWM speakers had more difficulty with initial nasal sounds in Standard Mandarin, which was similar to their performance in producing Standard Mandarin monosyllabic words. For English, production of (l) was significantly less accurate than (n), and (l) production in English was significantly worse than in Standard Mandarin. When both sounds occurred next to each other, there was a tendency toward producing only one sound, suggesting that the speakers assimilated production toward one phonological target. The results suggest that L1 influence may differ for the L2 and L3.

Keywords: /l/-/n/ merger, Southwestern Mandarin, English pronunciation, production, Standard Mandarin, L3 production, lateral nasal

INTRODUCTION

When the novel coronavirus that produces COVID-19 was first identified in the city of Wuhan, the Chinese government closed off the city of 11 million and a wider area with a population of 60 million. New hospitals were built, and doctors and other medical workers came from all over China to treat victims of the disease. One of the outcomes of this unprecedented lockdown was dialect/language contact between the Southwestern Mandarin (SWM) speakers of Wuhan and its related dialect areas and speakers of other dialects of Mandarin, resulting in the development of apps that translate the pronunciation of SWM speakers into Standard Mandarin for outside medical workers (Li et al., 2020). Without these technological solutions, there would have been widespread difficulty in communication, necessitating a larger number of interpreters who could bridge the gap between SWM patients and healthcare workers from other areas of China. Among the most challenging pronunciation differences for outsiders was the non-distinction between /l/ and /n/, a common marker of SWM.

The lack of an /l/-/n/ contrast in SWM, which is also found among Cantonese speakers (e.g., Hung, 2000; Ng, 2017) and in some other Mandarin-speaking areas of China such as Gansu province (e.g., Zhang, 2012; Han, 2014) is less well-known than some other highly studied pronunciation problems, especially the /l/-/ɹ/ contrast. Difficulty with /l/-/ɹ/ (e.g., lice-rice, collect-correct, fell-fair) is characteristic of Japanese English pronunciation, and it was heavily studied because of the ubiquity of Japanese learners of English starting in the mid 20th century. The mispronunciation of these sounds continues to stereotype Japanese English.

Both the /l/-/ɹ/ and /l/-/n/ contrasts involve the lateral /l/, and both involve another consonant similar to it in sonority (Yip, 2011), one a (coronal) approximant and the other a nasal. Phonologically, Japanese has neither /l/ nor /ɹ/, but rather has an apico-alveolar tap which shares some aspects of its articulation with both English /l/ and /ɹ/. Similarly, many SWM speakers appear not to have a contrast between the /l/ and /n/, which can cause challenges in being able to perceive and produce such a distinction in another language. Brown (1998) argues that when L2 speakers try to learn a phonemic contrast for which their L1 has no equivalent feature contrast (e.g., the coronal feature of English /l-r/ is not employed in Japanese), their perception of the novel contrast is unlikely to improve with increased proficiency. In contrast, learning a novel contrast when the L1 employs the same feature but not the same contrast leads to increased accuracy with increasing proficiency (e.g., English /l-ɹ/ for Chinese speakers who are familiar with the coronal feature from Chinese, or evidence that Japanese speakers improve on English /b-v/ because Japanese exploits the stop-continuant contrast in other sound contrasts). In regard to our study, it is possible that the /l/-/n/ merger in SWM speakers' phonological systems may lead them to produce their sounds with overlapping features of both /l/ and /n/. In other words, both alveolar laterals and alveolar nasals are featurally ambivalent and ambiguous. Mielke (2005) provides evidence that both sounds are ambiguous with regard to the feature (continuant) because in some languages laterals and nasals pattern similarly to (−continuant) sounds and in other languages they pattern similarly to (+continuant) sounds.

Mielke suggests that the feature (continuant) be underspecified for certain sounds (especially nasals and laterals) to account for the patterns evident in different languages. Similarly, Ladefoged and Maddieson (1996) demonstrate that laterals occur with multiple types of articulation, and that a posited feature (lateral) may or may not have a central closure. For nasals, Ladefoged and Maddieson (1996) document a variety of partially nasalized consonants, including nasalized approximants.

Besides featural differences, an inconsistent distinction between (l) and (n) may be the result of a laterally articulated nasal (Soejima et al., 1990; Yip, 2011), which would have aspects of both (l) and (n). Without an acoustic analysis of SWM speakers' productions, which is beyond the scope of this paper, it is not clear whether SWM (l) and (n) articulations can be described similarly to the sounds produced in English and Standard Mandarin. Unlike the more extensive research on (l) and (n) in Cantonese learners of English, there is little research on SWM speakers' pronunciation of (l) and (n) in either standard Mandarin or in English. The goal of this study is to provide a beginning to such research.

The study uses production data to describe the distribution of SWM speakers' production of /l/ and /n/ in two additional languages, Standard Mandarin and English. Because we have no equivalent perception data, we cannot make claims for why there are variations in the production of (l) and (n). Goto (1971) reminds us that adult learners may be more successful in producing sounds that they do not perceive, and that production data may overrepresent L2 learners' perception. It is also possible that the sound produced by SWM speakers may use a variable articulation that has both nasal and lateral elements (Soejima et al., 1990) that are heard by listeners with an /l/-/n/ contrast (such as the researchers) as fitting into one category better than another. Such variable articulations are common in languages. The /s/ in English, for example, may be articulated with the tongue tip up toward the alveolar ridge or the tongue tip down behind the bottom teeth and the tongue blade raised toward the alveolar ridge. This variation is rarely noticed by English speakers, but the tongue tip up articulation has been argued to be the cause of Japanese speakers' difficulty producing differences between /s/ and /ʃ/ in English words (Raver-Lampman et al., 2015).

SOUTHWESTERN MANDARIN

Southwestern Mandarin, the most widely used dialect of Mandarin (Li, 2009) with over 270,000,000 speakers (Chinese Academy of Social Sciences, 2012), is spoken mainly in nine provinces and regions including all of Sichuan, Chongqing, Guizhou, and Yunnan, and some areas in Hubei, Guangxi, Hunan, Shanxi, and Jiangxi (**Figure 1**).

Southwestern Mandarin was probably originally centered in Sichuan, then spread to Hubei, Guizhou, Yunnan, Guangxi and other provinces with Sichuan (including Chongqing municipality) as the center (Qian, 2010). Southwestern Mandarin has been classified into six dialect varieties, that is Chuanqian, Xishu, Chuanxi, Yunnan, Huguang and Guiliu. Each variety has several sub-varieties, with a total of 22 sub-varieties identified (Li, 2009; Qian, 2010; Chinese Academy of Social
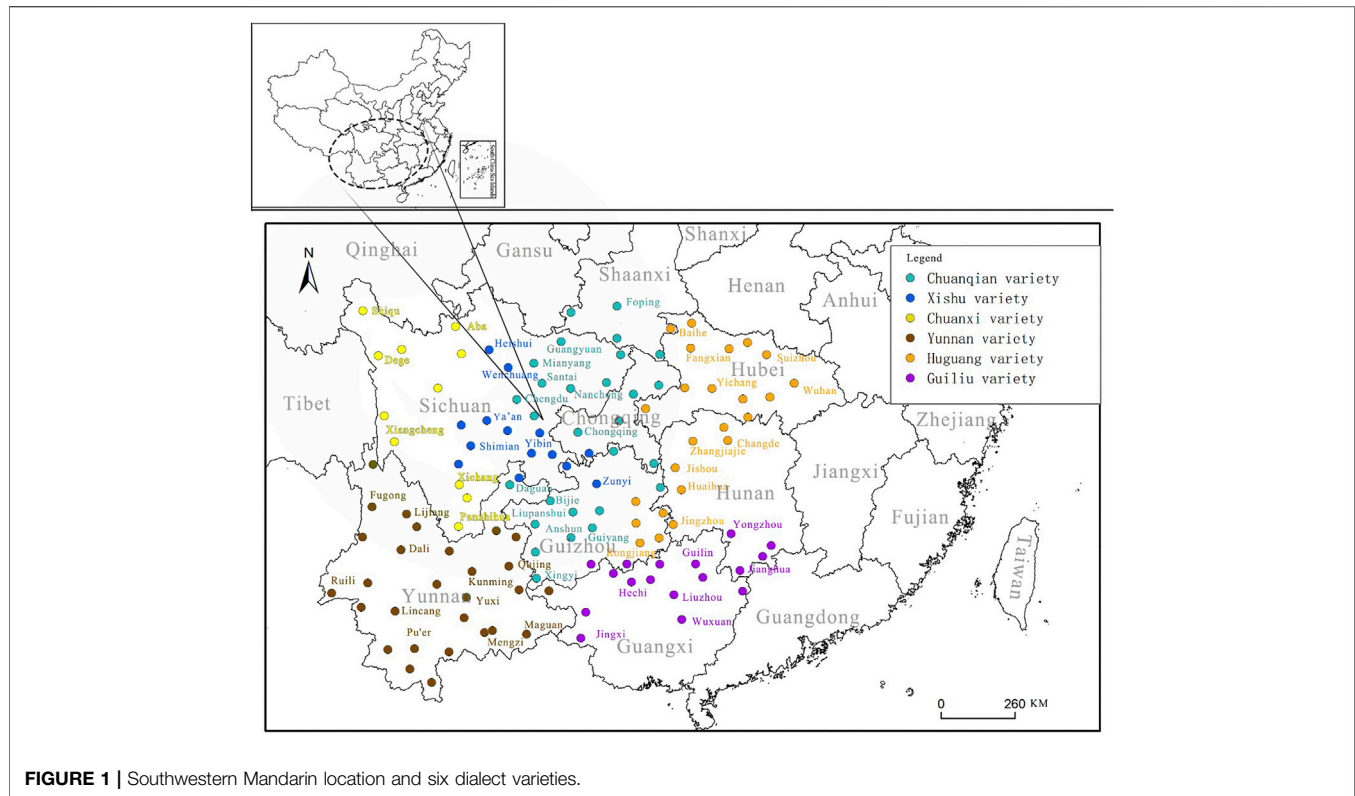
FIGURE 1 | Southwestern Mandarin location and six dialect varieties.

Sciences, 2012). Others have classified it into more dialects (Baker, 1993; Kupaska, 2010). Different dialect varieties have distinctive distributions of phonological features.

For this study, we look at the pronunciation of alveolar nasal and lateral consonants in SWM speakers' L2 and L3 production (Li, 2004; Sun, 2011). We first look at the production of all speakers to determine whether they represent different sub-dialects of SWM in their production of (l) and (n) (Qian, 2010). Like any large dialect area, Southwestern Mandarin has variation in how particular phonological features are realized. For example, the Chuanqian variety is reported to widely confuse /n/ and /l/. Only the alveolar nasal exists in the Chengyu sub-variety of the Chuanqian variety (Zeng, 2009; Zhou, 2014; Li, 2017), but in the Qianzhong sub-variety of Chuanqian, the opposite is true, with only a lateral evident but no nasal (Wang, 1981; Yuan, 1996). In the Xishu variety, the confusion of /n/-/l/ also exists, but the phonology of the sounds in the variety is not established (Luo, 2016). In contrast, the Yunnan variety is considered part of SWM because it shares other phonological features with other sub-varieties of SWM, but speakers of this variety appear to distinguish /l/ and /n/ (Wang, 1986; Li, 2004; Qian, 2010; Chen, 2013).

LITERATURE REVIEW

The /l/-/n/ Contrast in Cantonese and Southwest Mandarin

Nasal and lateral sounds are common across languages of the world, with only a few languages lacking nasals (Mielke, 2005).

Ladefoged and Maddieson (1996) also show that both nasals and laterals occur with multiple types of articulation (e.g., approximants, fricatives, clicks) and in different places of articulation (e.g., labial, velar, alveolar, palatal). Nasal consonants are produced by lowering the soft palate so air flows out through the nasal cavity, and the articulators completely stop the oral airflow. A lateral consonant is often but not always produced with a central obstruction, and air passes through one or both sides of the tongue (Ladefoged and Johnson, 2011). The two sounds thus differ mainly in the presence or absence of nasal airflow or in whether the air comes out of the nose or from the sides of the tongue. Acoustically, English (n) and (l) differ in their F2 (Koffi, 2019). In Standard Mandarin, the F2 of the Standard Mandarin (n) is often weak and sometimes even disappears, while the F2 of (l) sometimes shows only a low F2 value (Lin et al., 2013) but is higher when preceding a high vowel and lower when preceding a low vowel (Bao and Lin, 2014).

There are at least two Chinese languages in which /l/ and /n/ do not always contrast. The first is Southwestern Mandarin, and the second is Cantonese (Deterding, 2010). Of these two, Cantonese is far more extensively studied. Historically, Cantonese had both /l/ and /n/ phonemes, but the two sounds have merged over time toward becoming one phoneme with two allophones in complementary distribution (Zee, 1999): (l) occurs in syllable-initial environments, while (n) occurs syllable finally (Ng, 2017), a similar distribution to the light and dark /l/ sounds in English. As a result, Cantonese speakers do not produce (l) in coda position, and the dark (ł) found in English is commonly ignored altogether (Wong, 2008). Younger speakers in Hong

Kong have led this merger (Yeung, 1980; Tong and James, 1994; Chan and Li, 2000). Ng (2017) showed an incomplete merger of these two sounds. Initial /n/ was pronounced as (l) 72.3% of the time in Standard Mandarin words whereas initial /l/ was pronounced as (n) only about 4% of the time. In perception, Cantonese listeners misidentified (n) as (l) in 30.3% of items, while (l) was misidentified as (n) in 24.8% of items.

Because of the historical connections between Britain and Hong Kong, much of the research on the /l/-/n/ merger among Hong Kong Cantonese speakers has looked at the effect of the merger on their English pronunciation (e.g., Deterding, 2006; Deterding et al., 2008). Au (2011) looked at how HK Cantonese speakers produced initial /n/ and /l/ in English syllables. The research reported that words with initial (n) were produced invariably with (l). Chan (2014) does not mention /l/-/n/ as a problem for HK Cantonese EFL learners, indicating that initial /l/ is more likely to be confused with /ɹ/ or /w/, but this may be because Chan (2011), in a study of perception of English sounds by Cantonese EFL learners, found that their production was almost completely accurate (97%), and that learners were largely able to distinguish /l/ and /n/ initially (96% accurate). In an earlier contrastive analysis, Chan and Li (2000) characterized initial laterals in Cantonese in this way: initial (l) as in 'love' /lʌv/ may sometimes be pronounced with some "n" quality, giving the impression of a nasalized /l/ sound, viz. (l̃) (p. 80). Another study suggests that their perception may not be so accurate. Li and Hua (2014) studied the effects of visual cues on the ability of both groups to perceive /l/ and /n/ accurately. While they found that audiovisual input (audio supplemented by facial cues of native speakers) significantly helped Cantonese speakers to better distinguish the two sounds, the Mandarin speakers who served as a comparison group found the visual cues unnecessary because they already had a phonemic contrast between /l/ and /n/.

The non-final /l/ and /n/ of SWM speakers seems to follow different patterns from those attested for Cantonese, but there are many questions about /l/-/n/ in SWM that remain unanswered. [In coda position, SWM, like other Mandarin varieties, does not license the lateral but allows (n), although the realization of (n) in our data (Zhang and Levis, 2020) varied between a (+consonantal) pronunciation and a (−consonantal) sound, i.e., a nasalized vowel.] Although the historical loss of a distinction between these sounds in SWM is likely related to a historical process of denasalization (Soejima et al., 1990; Hu, 2007), different sub-dialects in SWM have different realizations of the historical change, as Soejima et al. (1990), p. 131 say:

> In some Chinese dialects, mainly in southern China, (including southern Mandarin dialects), the opposition between /n/ and /1/ has been lost which existed in so-called Ancient Chinese ... However, the reflexes of these phonemes in modern Chinese dialects and the environmental conditions of this merger are not uniform. That is, in some dialects these phonemes coalesced into /n/ (e.g., Changsha in Hunan, Chengdu in Sichuan), and in some other dialects into /1/ (e.g. Nanjing in Jiangsu). In some dialects, that coalescence occurred only in syllables without the medial front glide (j) (e.g. Nanchang in Jiangxi ...), while in some other dialects this coalescence occurred spontaneously.

In another possible example of a production that combines both lateral and nasal features Chan (1987) reports that younger Southern Min speakers produce (n) as (n̪l). Some sub-dialects of SWM are reported to have /l/ but not /n/ (Zhang, 2007 for Sichuan English), others /n/ but not /l/, and yet others have both phonemes yet fail to always contrast the sounds (Ao and Low, 2012). It is also clear that Cantonese and SWM pattern differently. In SWM, initial /n/ seems to be more stable than initial /l/ (Koffi, 2019), and there is little evidence of a complementary distribution in the two sounds, but neither do the sounds seem to be in free variation (Zhang, 2007). Ultimately, there is too little information about the distribution of the two sounds in the L1, and not enough is known about how the two sounds occur in L2 and L3 perception and production. Zhang (2007), for example, presented only a single case study of a speaker from Sichuan. Pennington and Saunders (2013) developed a pilot corpus of two speakers from Guiliu, a subdialect of SWM, in which they assert that "/l/ merges with /n/", yet without evidence for the assumption that /n/ is the phoneme that is more stable. Koffi (2019) provided an acoustic analysis of only five mispronunciations in three words from an online database. Ao and Low (2012), in a study of the English spoken in Yunnan province (part of the larger SW Mandarin dialect region) did not identify /l/ and /n/ as a likely problem for this sub-dialect, perhaps because Yunnan is sometimes reported to speak Northwestern Mandarin (Chan, 1987).

The /l/-/ɹ/ and /l/-/n/ Contrasts

We can hypothesize patterns of production of (l) and (n) from other research, especially the /l/-/ɹ/ contrast. There is strong research evidence from Japanese learners of English that learning to perceive and produce the /l/-/ɹ/ contrast is extremely challenging. A number of studies have documented significant improvements from training in both perception and production (Hazan et al., 2005; Aoyama et al., 2008), and there is also evidence that production accuracy can be better than perception accuracy (Goto, 1971), and that perception is a more reliable measure of whether an L2 speaker's production difficulties are based on fundamental difficulties in perceiving differences between two L2 sounds. Brown (1998) and Brown (2000) demonstrates that difficulties with nonnative phonemic contrasts are not all equivalent, proposing that when features in a learner's L1 sound system match features in the L2 (e.g., continuant), L2 phonological perception is likely to be more successful, but that when a feature important in distinguishing phonemes in the L2 system is not present in the L1 system [e.g., the non-use of the (coronal) feature in Japanese and Korean], L2 perception will be impaired regardless of increased L2 proficiency. Brown (2000) provides an example of the English /s/-/θ/ contrast, which because of the feature (distributed), was equally difficult to perceive for Japanese, Korean and Chinese learners of English whose L1s do not require a distinction between (distributed) and (anterior). Applied to the /l/-/n/

contrast, SWM may not have a featural distinction between (n) and (l) even though childhood classroom learning of Standard Mandarin could call attention to the contrast, attention that could be later reinforced by learning of English. Yip (2011), p. 12 argues that although there is conflicting evidence for a feature called (lateral), there is compelling indirect evidence for its reality:

> In Eastern Catalan (and Sanskrit), for example, (lateral) spreads onto nasals to create a lateral nasal: /nl/ → (l̃l) in /son les tres/ → (sol̃les tres) . . . There are well-known phonological processes that involve only (l) and (r), and in which they either dissimilate, as in Latin, where the suffix /-alis/ surfaces as (-aris) after a lateral root: nav-alis *vs.* milit-aris . . . or assimilate, as in Sundanese, where the infix /-ar-/ surfaces as (-al) after a preceding /l/: (k-ar-usut) *vs.* (l-al- əga).

In the absence of perception data, previous research on other contrasts offers hints about possible challenges faced by SWM speakers in producing a distinction between (l) and (n). Goto (1971) shows that Japanese speakers who cannot reliably distinguish /l/-/ɹ/ nevertheless may be able, because of varied spoken strategies, to produce the sounds correctly at a much higher rate. Brown (1998) explains that L2 learners often can "accurately produce a nonnative contrast even though the same learners are unable to distinguish the two sounds perceptually" (p. 156). This happens because they already have fully developed motor control, allowing them to often carry out necessary articulatory movements to be heard as they intend. For example, SWM speakers may be able to produce differences between (l) and (n) when they read words because they know that spelling differences in the written input reflect different sounds.

If the challenges in pronouncing this contrast come from SWM lacking a featural distinction between the two sounds, we would expect that production errors would be common in all environments, with some environments being more challenging than others. We would also expect that speakers from different varieties of SWM would have different frequencies of errors depending on whether the variety has been described as having an /n/, an /l/, or both phonemes (Qian, 2010). Ultimately, however, a production study cannot provide evidence for whether SWM speakers perceive the differences between the sounds when pronouncing an L2.

/l/-/n/, Functional Load and Speech Intelligibility

Many Chinese students study abroad in English-speaking countries, with around 370,000 in the United States alone (IIE, 2019). Of these, a substantial number likely come from the Southwestern Mandarin dialect area, which is also a destination for foreign students (Leung and Sharma, 2020). A non-distinction between /l/ and /n/ in English, although an identifying characteristic of their Chinese variety, may create difficulties in outsiders understanding their Chinese speech (as in the story at the beginning of this paper), but it is likely to be highly damaging to the intelligibility of their English speech. L2

pronunciation teachers who have not encountered SWM students may find themselves puzzled by this little discussed error (Richards, 2012). This can be illustrated by an experience of the second author. Once, a friend from Sichuan was going to go the supermarket to buy "wallets." When he was successfully helped by his housemate (after initial confusion and many corrections) to say "walnuts" instead, he immediately reverted to "wallets" thinking he was finally saying it correctly, but the production would likely create confusion in a supermarket.

In English, lack of a distinction between /l/ and /n/ is rated at the highest level of the functional load (FL) scale for onset consonants (Catford, 1987; Brown, 1988), a measure of how likely two contrasting sounds are to confuse listeners. (There is no equivalent measure for contrasts in Standard Mandarin, but in the view of the first author, /l/ and /n/ do not often contrast in Standard Mandarin.) High FL sound contrasts have many minimal pairs (as in low-know, light-night) and a correspondingly greater chance that mispronunciations will be heard as a different word. In comparison, low FL sound contrasts (such as thought-fought) have few minimal pairs, a lower likelihood of confusion for listeners, and a greater likelihood that listeners will be able to successfully understand an utterance even when there is a pronunciation error. Functional load for onsets is particularly important because such consonants are most likely to lead listeners to expect a particular cohort of words (Bent et al., 2007).

Functional load is a way to measure error gravity in regard to segmental mispronunciations. Errors at the segmental level are frequent because they are unavoidable in speech; they are also conspicuous signals of nonnativeness or dialect, and they correlate with how well speakers are understood (Zielinski, 2006; Munro et al., 2015). In one investigation of the effect of segmental accuracy on intelligibility judgments, Bent et al. (2007) found a strong correlation between judgments of intelligibility and segmental pronunciation accuracy of vowels and word-initial consonants. Zielinski (2008) similarly showed that errors in stressed vowels and consonants that occurred in stressed syllables were likely to damage intelligibility. These findings have one thing in common: each segment is not equally important for intelligibility, and thus teachers should prioritize teaching certain segments. This view finds its most complete explanation in the concept of FL (Catford, 1987; Brown, 1988; Sewell, 2017), which offers a reason why some segments are more important for intelligibility.

Although discussions of FL stretch back over 100 years, FL became widely noticed in the 1980s in regard to L2 pronunciation teaching (Catford, 1987; Brown, 1988; Catford, 1988). The FL model presents the relative importance of segments in terms of how much work two phonemes do in communicating meaning differences in a language (Sewell, 2017). When words differing in one sound (i.e., minimal pairs) are more frequent, the corresponding sound contrast has a higher functional load in the language. More complex measures of FL can be made by taking various criteria into account, such as the number of minimal pairs that a particular phonemic distinction differentiates at the beginning of a word and end of a word (Catford, 1987) and the frequency of occurrence of each word in a

minimal pair; pairs of words that are both frequently encountered have higher FL than those that are infrequent. Part of speech in a minimal pair is another contributing factor. There is a higher value when two words share the same part of speech, as listeners are thought not to confound words as easily if the words are, for example, a noun and a pronoun (Brown, 1988). Brown (1988) also presents a hierarchy of the English phonemic contrasts ranging from 1 (low) to 10 (high). There is empirical evidence to support the use of FL in predicting judgments of comprehensibility and accentedness. Munro and Derwing (2006), using read-aloud sentences, explored the difference between high FL and low FL consonantal errors (including /l/-/n/) on ratings of comprehensibility and accentedness. They found that high FL errors had a significantly larger effect on judgments of comprehensibility and accentedness, and they also found an effect of frequency for ratings of accentedness, that is, two high FL errors caused greater severity in judgments of accentedness, but this was not true for multiple low FL errors.

Kang and Moran (2014) correlated high and low FL errors with test scores from the Cambridge ESOL General English examinations. Low proficiency scores correlated with more FL errors, while higher scores did not have many high FL errors. Low FL errors, in contrast, though more frequent than high FL errors, were similar in frequency for most levels. Suzukida and Saito (2019), in another study looking at the correlation of high FL errors with scores on two tasks with Japanese English learners in Canada, found that segmental errors with high FL values were more detrimental to judgments of comprehensibility than were low FL errors.

Because the /l/-/n/ contrast is likely to affect listener judgments of how comprehensible SWM speakers are and because it has been insufficiently studied in previous research, we want to describe the patterns of pronunciation for initial and medial /l/ and /n/ as well as to look at how speakers pronounce the two sounds when they occur together in the same word (e.g., walnut, only).

L2 and L3 Phonological Acquisition
This study looks at the effects of L1 on /l/-/n/ production in two additional languages, Standard Mandarin and English. For SWM speakers, Standard Mandarin is learned earlier (from the beginning of formal schooling) than English, and Standard Mandarin is typologically similar while English is not. There is conflicting evidence about the effects of the L1 or L2 on L3 pronunciation. Hammarberg (2001) says that "there appears to be a general tendency to activate an earlier secondary language in L3 performance rather than L1" (p. 23). The L3 can also influence the L2, as Cabrelli Amaro (2017) points out that "an ostensibly native-like L2 is more vulnerable to L3 influence than an L1" (p. 699). Llama et al. (2010) point out that the two primary factors involved in cross-linguistic influence for L3 acquisition are language distance (i.e., whether they two languages are typologically similar or dissimilar) and the language status of the L2—that is, the proficiency strength of the L2. Llama et al. (2010), who studied the voice onset time (VOT) of French-English and English-French bilinguals learning Spanish as an L3, found that the L2 was the stronger influence on VOT production of L3 Spanish. In another study, however,

Wrembel (2013) found that Polish L1 speakers (with French as their L2 and English as their L3) showed the strongest influence of their L1 on L3 accentedness. Their L2 French, in contrast, showed much less Polish influence. She suggests that L2 influence is likely to be strongest in the earliest stages of proficiency. She also points out that even though their French L2 was more advanced, the participants had learned English as a second L2 but were now coming back to it after some years. As a result, the stronger influence of language distance may have occurred because both French and English were simultaneous L2s.

Our study is not able to state unambiguously whether the L1 or L2 is the stronger influence on /l/ and /n/ production because we have no data on the participants' L1 production. Cabrelli Amaro (2012) makes clear that the features being examined must be represented by equivalent data from all three languages. However, Standard Mandarin is both typologically closer and likely to be more advanced in proficiency because of earlier learning. This would indicate that accurate pronunciation of the English /l/-/n/ contrast, if influenced by the L2, would have a similar error rate. However, if the error rate is much larger than that in the L2, it would suggest that L1 influence is stronger.

This Study
Our exploratory study examines how common /l/ and /n/ pronunciation deviations are in the L2 speech of SWM speakers. Although /l/ and /n/ substitutions are noticeable in SWM L2 speech, and those substitutions are likely to affect intelligibility, SWM speech is understudied in comparison to /l/-/n/ in Cantonese L2 speech or /l/-/ɹ/ in Japanese L2 speech. Because of this, our study describes what kinds of mispronunciations are likely in SWM speakers' production in other languages, whether their L2 pronunciations can shed light on their L2 phonology, and what information it provides for L2 pronunciation teaching. We look at five research questions that focus on the effect of the L1 phonological system on errors, error frequency for /l/ and /n/ production in both Standard Mandarin and English, and the influence of linguistic environment on production accuracy.

1) In regard to the production of /l/ and /n/, is SWM a consistent dialect, that is, do all subjects have difficulty producing a distinction between the two sounds?

   Because there is evidence that different dialects of SWM have differing patterns of (l) and (n) production, we first established whether all participants have difficulty in producing the distinctions. Those who had no difficulty in producing the target distinction would be excluded from further analysis so as to provide more accurate results for those who did not distinguish the two sounds.

2) Does (l) and (n) production vary by Tone and Rhyme in Standard Mandarin? Does accuracy change when (l) and (n) are pronounced in Standard Mandarin disyllabic words?

TABLE 1 | Subjects' dialect variety and hometowns.

Dialect variety	City
Subjects whose dialect/city is attested as having /l/ but not /n/ (*n* = 14)	
Xishu	Yibin; Zunyi
Huguang	Jishou; Changde
Chuanqian	Anshun; Bijie; Liupanshui; Guiyang; Xingyi
Subjects whose dialect/city is attested as having /n/ but not /l/ (*n* = 4)	
Chuanqian	Mianyang; Nanchong
Subjects whose dialect/city is attested as having /l/ and /n/ (*n* = 7)	
Yunnan	Qujing; Kunming; Dali; Mengzi; Hechi

The purpose of this question is to explore whether linguistic environment affects accuracy of /l/-/n/ production in Standard Mandarin reading. Environment here involves Tone and Rhyme (especially the vowel following the /l/ or /n/).

3) Is there an effect of linguistic environment on the accuracy of production in English words? To what extent do subjects mispronounce initial (n) or (l) in English both when the sounds occur alone and when there are competitor sounds at the end of the word?

This question addresses whether the presence of a competitor sound in the coda of the word (lemon, label, nine, nail) as opposed to the target sound being alone in the word (light, night). If participants make more errors in production with a competitor sound, this would suggest that the accuracy of initial (l) and (n) are affected by the presence of a similar sound.

4) Do SWM speakers produce /n/ and /l/ significantly differently in Standard Mandarin than they do in English?

This question asks whether the L1 shows different effects on the production of (l) and (n) in the L2 (Standard Mandarin) and L3 (English). It is possible for the L1 to similarly affect both the L2 and L3, or for the L2 to affect the L3 more than the L1. If L3 English has a higher rate of production errors than L2 Standard Mandarin, this would suggest that the L1 has a different impact on the additional languages, and that the L2 does not impact the L3.

5) When /l/ and /n/ occur word medially in English, are there differences in accuracy when the sounds occur alone or in combination with the other sound?

This question looks at potentially dissimilar phenomena. Medial (l) and (n) productions occur singly, in this case with words with orthographic

doubling in which the orthography represents a single sound, or together, that is, when both sounds are pronounced (medial <ln> and <nl>). We expect that words with both sounds (walnut, only) will show a higher error rate while words with orthographically doubled letters (swallow, winner) will have error rates similar to initial (l) and (n).

METHOD

Subjects

Twenty-five native speakers of Southwestern Mandarin, all current undergraduate students of Qufu Normal University, China, were recruited in a manner consistent with the institutional research guidelines of the university. All were compensated for taking part in the study. All grew up within the SWM area before attending college outside of the SWM area. A questionnaire was used to obtain their biographical information. All subjects indicated SW Mandarin was their first oral language, and it was also their dominant language for daily communication. Their hometowns and dialect areas are listed in **Table 1**. The subjects were divided into three groups according to their SWM dialect varieties and their first language phonological system: the /l/ group were subjects whose dialect is attested as having /l/ but not /n/ (Wang, 1981; Xiao, 1996; Ming, 1997, 2005; Zheng, 1999; Li, 2002; Wang et al., 2009; Xin, 2013; Luo, 2016); the /n/ group were those whose dialect was attested as having /n/ but not /l/ (Zeng, 2009; Zhou, 2014; Li, 2017) and the /l/-/n/ group were those whose dialect is reported as having /l/ and /n/ (Wang, 1986; Lan, 1995; Su, 2010; Chen, 2013; Wei, 2018). Note that all varieties except the Chunqiang variety are connected to a single pattern of /l/-/n/ production. All subjects learned Standard Mandarin and English when they started their formal education. They started to learn Standard Mandarin when they were in kindergarten at the age of 4 or 5 years old, and to learn English at the third year of primary school (around 9 years old). Their mean age was 19.5 years (range 18–21), their mean length of formal English instruction was 10 years (range 7–12 years), all of which was conducted in China (English is a core subject in the curriculum of elementary, secondary and tertiary education in China). They were not majoring in language and had not had formal phonetic training, although phonetic symbols are part of normal English instruction. All reported normal speaking and hearing abilities.

Stimuli

Subjects read aloud a word list which included English and Standard Mandarin words. The English word list included words with /n/ and /l/ in onset, medial and coda positions (**Table 2**). (Codas were included to examine the effect of competitor sounds within the same word, but they were not analyzed for this study.) For initial /n/, there were words with only initial /n/ (night), words with initial and final /n/ (nation), and words with initial /n/ and final /l/ (nail). Words with initial /l/ followed the same pattern. Example words include light, label, and lemon. The three different types of words for initial /l/ and

TABLE 2 | English word list numbers for initial and middle /l/ and /n/ recording.

Word type (n)	Words/Total	Word type (l)	Words/Total
Initial (n)		**Initial (l)**	
#n (night)	15/372	#l (light)	15/375
#n___n# (nation)	9/225	#l__l# (label)	9/225
#n___l# (nail)	10/246	#l___n# (lemon)	10/249
Medial (n)		**Medial (l)**	
#__nl__# (only)	10/250	#__ln__#(walnut)	7/174
#__<nn>__# (winner)	6/150	#__<ll>__#(swallow)	6/148

/n/ were used to explore whether the presence of a competitor sound at the end of the word would affect the accuracy of initial /l/ and /n/ production. Evidence from L1 acquisition shows that children may assimilate initial consonants to the articulation of following consonants (e.g., dog pronounced as *gog, see Pater and Werle, 2003). This led us to include words with word-final /n/ and /l/ to determine whether these expected sounds influenced accuracy of initial /l/ and /n/.

The words with /l/ and /n/ in medial environments included words with <nl> or <ln> (only or walnut), words with <ll> (swallow), and words with <nn> (winner). The list included 97 English words two times each, 34 with word-initial /n/ and 34 with word-final /l/, 10 with word-medial /nl/, seven with word-medial /ln/, and six each with word-medial /l/ and word-medial /n/. There were no distractor items. All English words were presented using both normal orthography and phonetic symbols based on British phonetic symbols used in Chinese EFL textbooks. (Chinese English learners start to learn English phonetic symbols when they are in junior high, and it was our expectation that using phonetic symbols could help them to understand the pronunciation of the words.)

Also included in the reading task were 50 Standard Mandarin monosyllabic words, all spoken with each of the four tones. All Standard Mandarin words were presented in Pinyin, which is the official romanization system for Chinese in China. In some cases, there was a gap and no actual Standard Mandarin word existed, resulting in a nonsense word, and subjects were told that some of the words they read would not be real words. They were nonetheless able to complete the task without undue difficulty. The total tokens were therefore 400, included 200 initial nasal words and 200 initial lateral words. Subjects read each word two times; the total analyzed initial /l/ tokens were 5,000 (25 initial l-words × 4 tones × 25 subjects × 2 times), with the same number for initial /n/ tokens.

The syllable structure of SW Mandarin is the same as Standard Mandarin and maximally consists of onset (consonant), (glide) and rhyme (with an optional nasal coda), as specified in Třísková (2011). For most words, the effective syllable structure is CV, far simpler than English. There are also four tones in Standard

Mandarin, i.e., Tone 1 (level), Tone 2 (rising), Tone 3 (falling-rising) and Tone 4 (falling). Rhymes include vowel (plus glide) with an optional final nasal. Standard Mandarin has a Sihu rhyming system with four kinds of rhymes, called Kaikou Hu, Hekou Hu, Qichi Hu and Cuokou Hu. Linguistically, Kaikou Hu is the rhyme that begins with non-high vowels, that is, not /i u y/, Hekou Hu begins with the /u/ sound, Qichi Hu begins with /i/, and Cuokou Hu begins with /y/. The Standard Mandarin words we used in our study were formed with an initial nasal or lateral with each of the four kinds of rhyme. According to the Mandarin Phonetic Alignment Chart Huang and Liao (2011), initial nasals and laterals can be followed by all four Sihu rhymes, but not with all vowels in each rhyme. In our words, we used only phonotactically-licensed rhymes, even though these did not always create an actual word in Standard Mandarin (**Table 3**). In the first word type, the vowels in pinyin were <a> (a), <e> (ɤ), <o> (o) or rhyming units that began with these vowels, such as <ai>, <ei>, <ou>, etc. In the second word type, the vowels were the high front vowel (i) or rhyming units beginning with <i>, such as <ie> (iɛ), <ing> (iŋ), etc. The third word type preceded the high back vowel (u) or rhyming units beginning with (u), such as <uo> (uo), <uan> (uan), etc. In the last word type, /n/ or /l/ preceded the high front vowel (y) or rhyming units starting with <ü>, like <üe> (yɛ).

Chinese words can be divided into monosyllabic words, disyllabic words (two monosyllabic words that function as a single word), trisyllabic words and tetrasyllabic words according to the number of syllables, and the syllables that make up a word are generally non-variable (Modern Chinese Teaching and Research Program, Chinese Department, Peking University, 2020). We used CNCORPUS (www.cncorpus.org) to identify whether the Chinese disyllabic stimuli were a (phonological) word or a phrase. Standard Mandarin disyllabic words have the features of word stress (Duanmu, 2000; Feng, 2016; Zhou, 2018; Zhang, 2021). Of the sixty disyllabic words, 56 were considered to be phonological words rather than phrases. The sixty Standard Mandarin disyllabic words with nasals and laterals in combination with other nasals and laterals (combo-type) or alone (non-combo type) were also part of the reading. Non-combo disyllabic words had one word with an initial nasal or lateral, while the other word did not, like nan fang (nan faŋ). In combo-type disyllabic words, both words had an initial nasal or lateral, such as nie lian (niɛ lɛn) (**Table 4**). The total recorded tokens for non-combo Standard Mandarin disyllabic words were 1,000, and for combo 2,000. In combo-type words, the second sound was a competitor (as in the words like lemon and label for English). In non-combo type words, only the word with the targeted sound was analyzed. Standard Mandarin disyllabic words were presented in simplified Chinese script, which is the norm in China. The Standard Mandarin reading items were also presented without distractors.

Procedures

The recordings were collected in a professional recording studio at Qufu Normal University. The productions were recorded to a computer *via* a directional microphone using Audacity software,

TABLE 3 | Standard Mandarin word list with initial /l/ and /n/.

Type	Word
Non-high vowels rhymes	na, la, ne, le, nai, lai, nei, lei, nao, lao, nou, lou, nan, lan, nang, lang, neng, leng, nong, long
/i/ rhyme	ni, li, nia, lia, nie, lie, niao, liao, niu, liu, nian, lian, nin, lin, ning, ling, niang, liang
/u/ rhyme	nu, lu, nuo, luo, nun, lun, nuan, luan
/y/ (ü) rhyme	nü, lü, nüe, lüe

Note: <ü> represents (y).

TABLE 4 | Standard Mandarin disyllabic words list numbers for initial /l/ and /n/ recordings.

Combo phrase type (n)	Words/Total	Combo phrase type (l)	Words/Total
#n+X (nan fang)	5/125	#l+X (lan fang)	5/125
X+#n (huang ni)	5/125	X+ #l (huang li)	5/125
#n+#n (niu nai)	10/250	#l+#l (liu lliang)	10/250
#n+#l (nie lian)	10/250	#l+#n (liu nian)	10/250

Note: "X" refers to an onset consonant that is neither (n) nor (l). These are classified as non-combo type phrases.

with a sampling rate set to 44.1 kHz at 16 bits per sample on one channel. Before recording, all speakers were given sufficient time to practice all words. Speakers were asked to read each word two times. The duration of the reading was about 15–20 min for both English and Mandarin. The mono recordings were saved as individual WAV files for evaluation. Recordings that were problematic (e.g., from subject noise or from sitting too far from the microphone) were excluded from the analysis.

Subjects first read the Standard Mandarin monosyllabic words, each spoken with all of the four tones, followed by the sixty Standard Mandarin disyllabic words. Finally, subjects also read all the English words two times each. Certain tokens were excluded because three subjects missed some words while reading or because subjects sat too far from the microphone or made other noises.

Data Analysis
All English sound files were evaluated by the researchers, and all Standard Mandarin sound files were evaluated by the first author, a native speaker of Standard Mandarin. Both researchers in this study have linguistic training and are trained in listening to second language pronunciation learners. One researcher is a native speaker of American English with 35 years of experience. The other researcher is a native speaker of Standard Mandarin and is an advanced L2 speaker of English who teaches English in China. Errors in the production of initial /l/ and /n/ were identified as a substitution. We evaluated the second production of each pair. This allowed subjects to read each word once before producing words that were evaluated. As a result, if the first reading was correct and the second one was incorrect, we rated it as incorrect. If the first was incorrect and the second was correct, we rated it as correct. If the researchers were not in agreement, we consulted Praat, especially looking at the F2 measures (Koffi, 2019) and discussed our decisions until we agreed.

RESULTS
The first research question asked whether our subjects all had difficulty producing the distinction between (l) and (n). **Figure 2** shows the percentage of errors from the original 25 SWM speakers in this study. Seven, all from the Yunnan variety, were found to have no difficulty producing a difference between the two sounds in Standard Mandarin, in accord with some descriptions of Yunnan speakers of English (Deterding, 2006; Ao and Low, 2012). Their mean error rates for (n) production and (l) production were 0.009 and 0.014%, respectively. As a result, we report results only from the remaining 18 subjects, 14 from the /l/ group and four from the /n/ group. Because of the disparity in numbers of subjects, any comparisons between these two varieties can only be reported descriptively. It appears that the /n/ group had a greater difficulty with the production of (l), as would be expected, and that the /l/ group's errors for (n) and (l) had more errors for (n) readings. Mean error rates for the /l/ group's production of (n) and (l) were 14.61 and 7.25%, respectively, while the /n/ group's mean error rates for production of (n) and (l) were 7.63 and 16.25%, respectively. These proportions are mirror images of each other and indicate an influence of the L1 phonological systems.

Mean Error Rate of Initial /l/ and /n/ in Standard Mandarin
Figure 3 shows the mean error rate of initial /n/ production for the four types of rhymes in Mandarin. SWM speakers had the largest number of errors for initial nasals in Standard Mandarin (25%) when the rhyme started with /y/ and with /u/ (23.05%). The other rhymes, in descending order, were rhymes starting with non-high vowels and with /i/ at 20.98 and 18.05%, respectively. When SWM speakers pronounced initial laterals in Standard Mandarin, the rhymes may have influenced the accuracy of their production, but the variations were not as noticeable. In descending order, the mean error rates for initial /l/ were at 15.74% for Kaikou Hu (non-high vowels rhyme), 18.05% for Hekou Hu (/u/-rhyme), 21.45% for Qichi Hu (/i/-rhyme), and 16.66% for Cukou Hu (/y/-rhyme) (**Figure 3**). These results suggest that SWM speakers may be generally more accurate with initial nasals and laterals when the rhymes start with a nuclear non-high vowel.

We also examined SWM speakers' performance on disyllabic words in Standard Mandarin. The results indicate that SWM speakers had more difficulty in producing initial nasal sounds
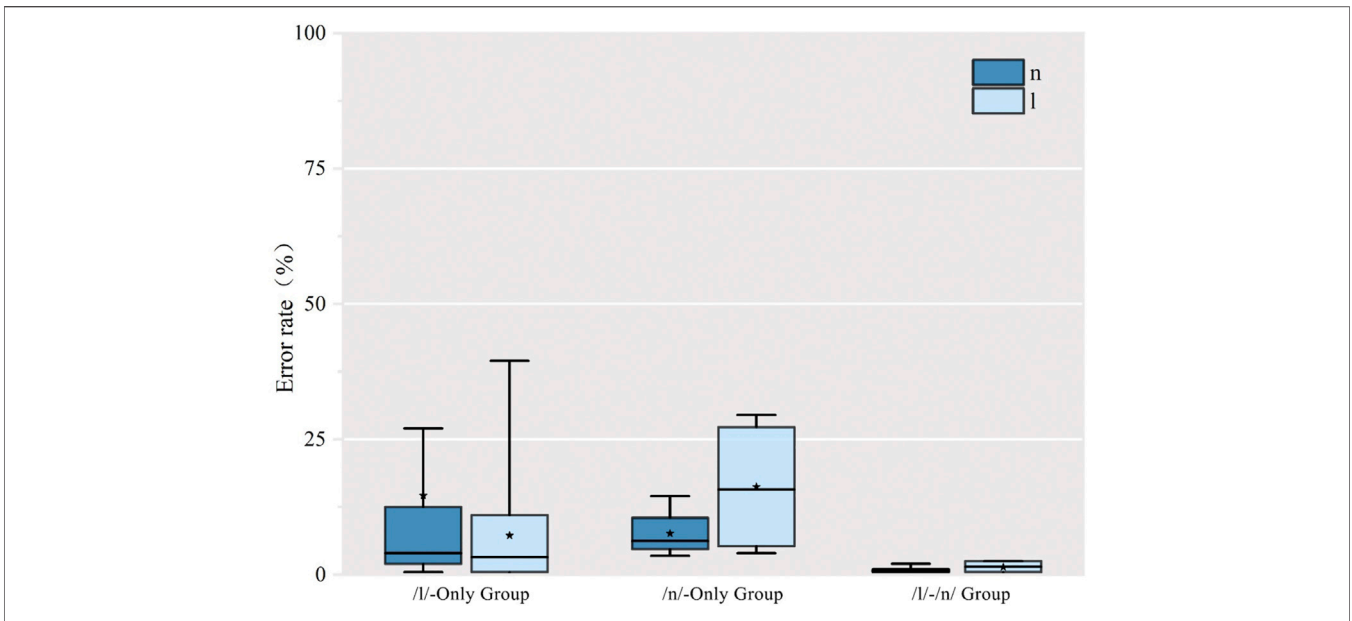
FIGURE 2 | Error rate of production of Standard Mandarin words with initial /n/ and /l/ by subjects' variety.
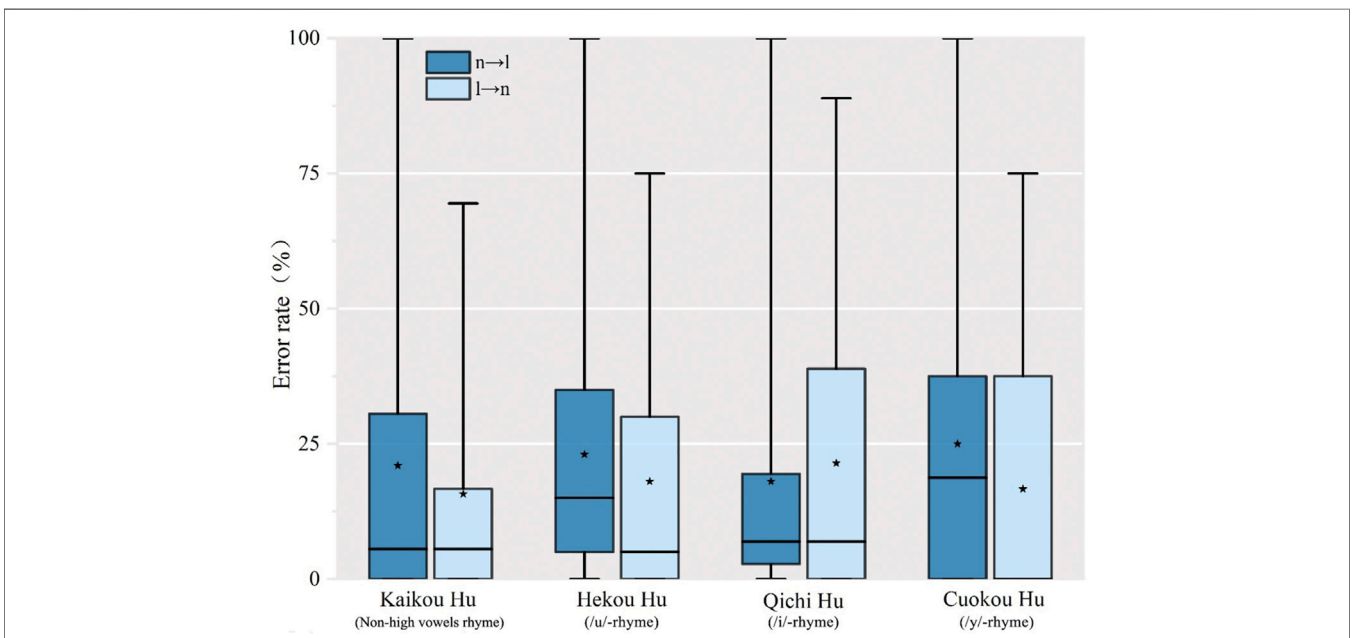

FIGURE 3 | Error rate of production of Standard Mandarin words with initial /n/ and /l/ by word types.

than lateral sounds in Standard Mandarin (see **Figure 4**), which was similar to their performance in producing Standard Mandarin monosyllabic words. There were two types of Standard Mandarin disyllabic words analyzed. **Figure 5** shows SWM speakers had more trouble producing initial nasal and lateral sounds in non-combo Standard Mandarin disyllabic words.

The mean error rate for producing initial nasals in non-combo disyllabic words in Standard Mandarin was 22.77%, and the mean error rate for initial laterals in non-combo disyllabic words was 15.55%. However, the mean error rate of production for combo disyllabic words was 18.05% for initial nasals and only 7.5% for initial laterals, which indicates, somewhat surprisingly, that the
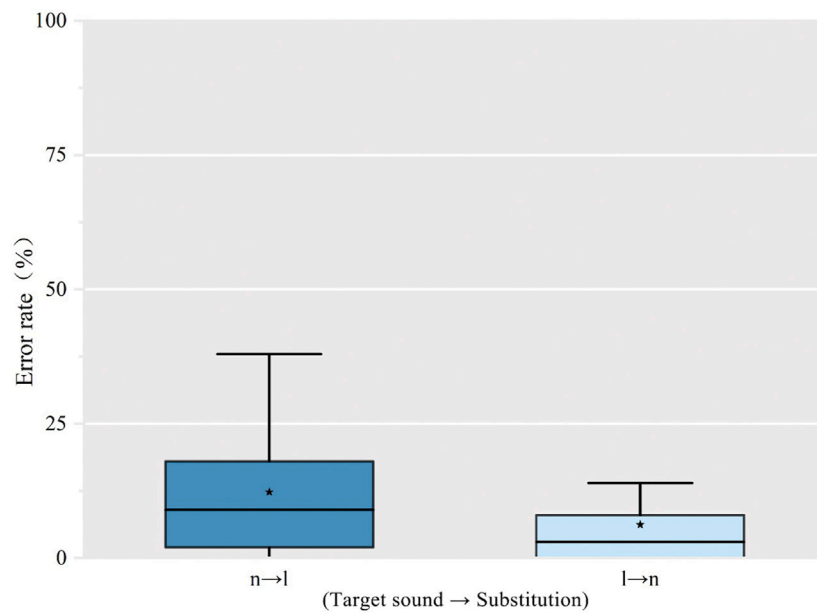
FIGURE 4 | Error rate of production of Standard Mandarin disyllabic words with initial /n/ and /l/.
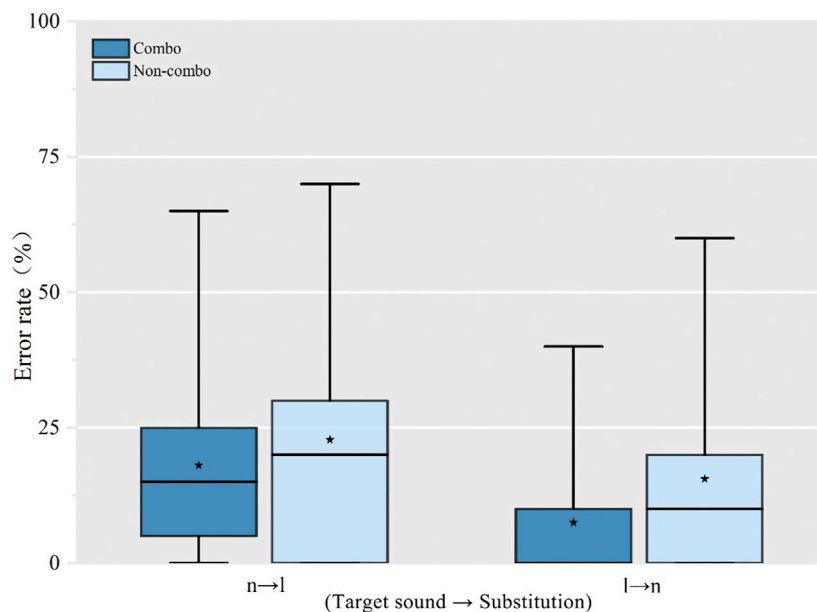

FIGURE 5 | Error rate of production of Standard Mandarin disyllabic words by formation.

presence of a competitor sound facilitated accuracy for both initial nasals and laterals. Rather than making the articulation more challenging, subjects appeared to do better. When the competitor sound was identical, this makes sense as both can more easily assimilate to the same articulation. When the competitor was different (as in lemon, nail), this greater accuracy suggested greater care in distinguishing the sounds.

A mixed-effects logistic regression analysis using lme4 in R was used to examine the effect of tone on production of initial nasals and laterals in Standard Mandarin (R Development Core

TABLE 5 | Descriptive results for tone effects.

		β	SE	z	p
NL1	(Intercept)	2.936	0.800	3.673	0.000
	Tone 2	−1.027	0.881	−1.166	0.243
	Tone 3	−0.766	0.908	−0.844	0.398
	Tone 4	−1.248	0.863	−1.446	0.148
Tone 1	NL2	0.139	0.528	0.263	0.792

Notes: NL1 = lateral, NL2 = nasal.

TABLE 6 | Tone comparison results under the condition of lateral.

Tone pairwise	β	SE	z	p
1–2	1.203	1.210	0.994	0.320
1–3	0.732	1.274	0.575	0.565
1–4	1.221	1.208	1.01	0.312
2–3	−0.47	0.998	−0.471	0.637
2–4	0.017	0.911	0.019	0.985
3–4	0.488	0.995	0.490	0.624

Notes: 1 = tone 1,2 = tone 2, 3 = tone 3, 4 = tone 4

TABLE 7 | Tone comparison results under the condition of nasal.

Tone pairwise	β	SE	z	p
1–2	0.801	1.288	0.622	0.534
1–3	0.789	1.290	0.612	0.540
1–4	1.267	1.227	1.033	0.301
2–3	−0.012	1.069	−0.011	0.991
2–4	0.466	0.991	0.470	0.638
3–4	0.478	0.994	0.481	0.631

Notes: 1 = tone 1,2 = tone 2, 3 = tone 3, 4 = tone 4

Team, 2009) with four tones, and two consonants (nasal and lateral) as fixed factors, subject and item number as random factors. The results (**Tables 5–7**, **Figure 6**) showed that there was no significant difference between subjects' error rates for lateral and nasal production for the four tones. Tone did not affect mispronunciations of initial /l/ or /n/ for SWM speakers.

Mean Error Rate of Initial and Medial /l/ and /n/ Production in English

Figure 7 shows that SWM speakers had consistent difficulties in the production of syllable initial /l/ with all three types of initial /l/ English words. Descriptively, the mean error rate for each was high, with 30% or more productions being wrong for each environment. Words with initial lateral and final alveolar nasal (lemon) were mispronounced 37.77% of the time, followed by the words with initial and final lateral (label) at 35.18%. The lowest mean error rate was for words with only an initial lateral (light), which was 31.11%. SWM speakers had much less trouble pronouncing nasals. The mean error rate was around 15% with modest variation between environments. The mean for the words only with initial nasal (night), the words with initial nasal and final lateral (nail), and initial and final nasals (nation) were 21.48, 15 and 14.51%, respectively. These descriptive results suggest that environments may affect initial /n/ and /l/ differently. While the absence of a competitor sound meant worse accuracy for /l/, it meant better accuracy for initial /n/.

English is the third language for all subjects; we also wanted to know how different groups of SWM speakers pronounced English words with initial /l/ or /n/. As with Standard Mandarin, the effect of the participants' sub-variety is
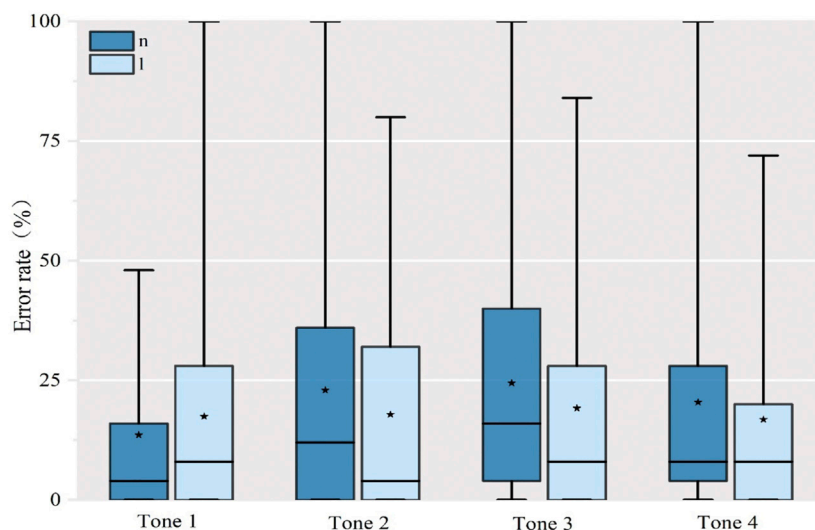

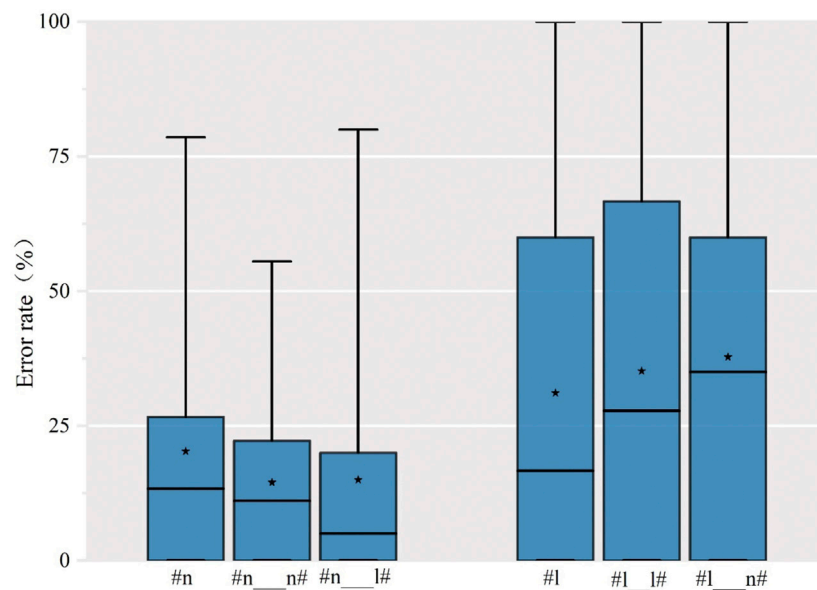**FIGURE 6 |** Error rate of production of Standard Mandarin words by tone.

FIGURE 7 | Error rate of production for the syllable initial /n/ and /l/ in English words by word types.

evident (**Figure 8**). For both the /l/-only group and /n/-only group initial (n) was much more accurately produced than initial (l), suggesting that even for speakers for whom initial /n/ is not phonemic, the sound is more likely to be pronounced accurately. This may be because most Mandarin varieties license nasals as codas. In other words, the feature (nasal) exists in Mandarin even when it is not phonemic in a particular environment, and the /l/-only group is therefore as successful as the /n/-only group in using this feature in their L2 and L3

nasal production. It is in the production of initial /l/ that we see differences, with the /n/ group being much less accurate than the /l/ group. In light and lemon type words, the /n/ group was far less accurate, but in label type words, the /l/ and /n/ groups showed similar accuracy. This may be because words with initial and final /l/ helped them to focus on the production of initial /l/, whereas /l/ words with final /n/ resulted in a much higher mean error rate for the /n/ group, suggesting that the presence of both sounds made it more challenging for them to
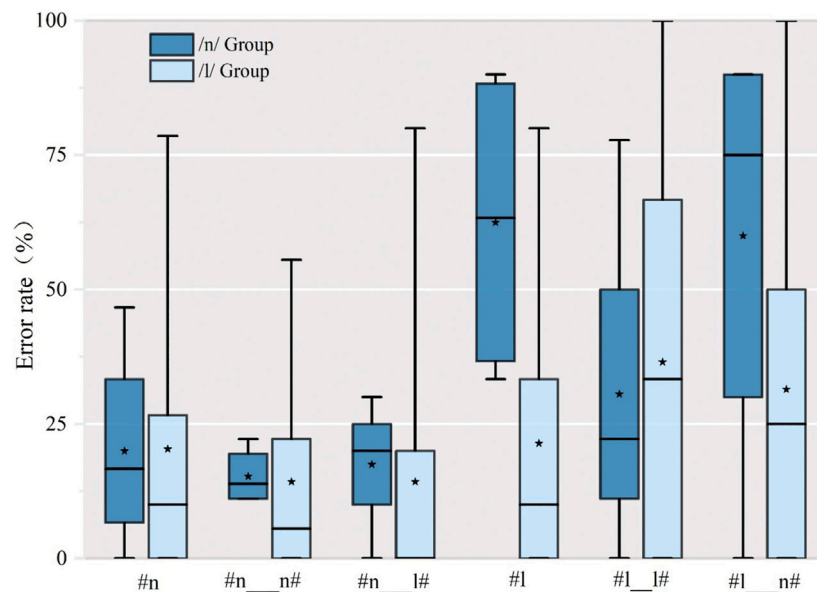

FIGURE 8 | Error rate of production for the syllable initial /n/ and /l/ in English words by subjects' variety.
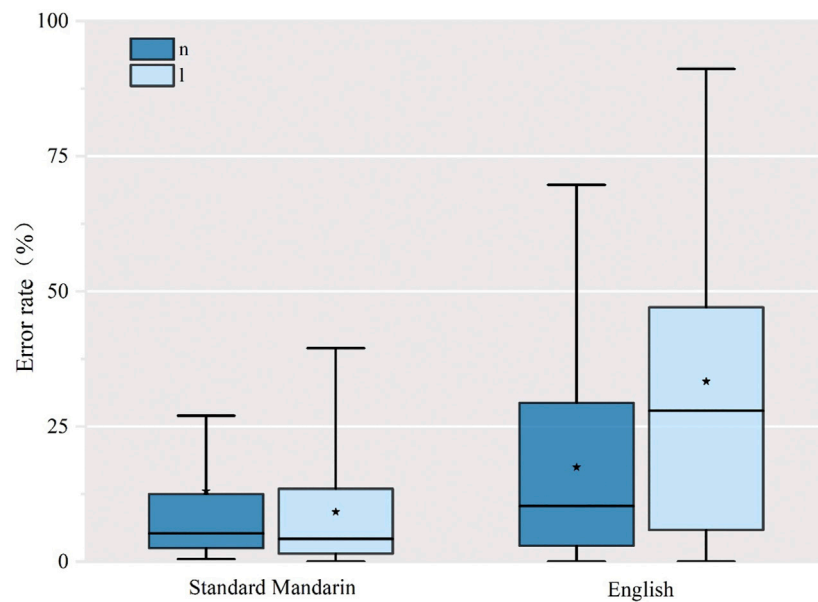
FIGURE 9 | Error rate of production of words with initial /n/ and /l/ in Standard Mandarin and English.

articulate the competing sounds in the same word. The /n/ group had a similar pattern for initial /n/ words, with n_l words being more difficult than n_n words, albeit with a smaller difference.

Influence of L1 on L2 (Standard Mandarin) and (L3) English Production

SWM speakers demonstrate production problems with the initial /l/ and /n/ when they speak in both Standard Mandarin and English (see **Figure 9**). In Standard Mandarin, the mean error rate for initial /n/ was slightly higher than for initial /l/. In English, their performance was different. The mean error rate for syllable initial /l/ in English (33.37%) was much higher than in the syllable initial /n/ (17.49%), with greater variability in performance.

A 2 × 2 within-subjects ANOVA with language (English and Standard Mandarin) and consonants /n l/ as factors showed a significant interaction [$F(1, 17) = 11.916$, $p = 0.003$]. Simple effects indicated that English productions had a greater number of errors than Standard Mandarin both in nasal and lateral sounds. For lateral sounds, the errors in English were significantly higher than those in Standard Mandarin (M =0 .238, SD =0.049, p =0.000). There was no significant difference between production of the nasal in Standard Mandarin and English.

We computed the Spearman correlation of SWM speakers' production of Mandarin and English. The correlation showed that the accuracy of initial nasal production for English was highly correlated to the initial nasal production of Standard Mandarin (r =0.898, p <0.001) as was the production of initial lateral sounds in Standard Mandarin and English (r =0.778, p <0.001). This means that if our SWM subjects had more

accurate pronunciation for initial nasals and laterals in Standard Mandarin, they were likely to also be more accurate in English.

Medial (l) and (n) in English, Alone and in Combination

In English, SWM speakers demonstrated serious difficulty when both nasal and lateral sounds occurred together in medial position (**Figure 10**), with #_ln_# words mispronounced at a higher rate than #_nl_# These words were mispronounced in a large majority of all productions. SWM speakers consistently produced only a single nasal or lateral rather than two distinct sounds. It appears that the /n/-only group, small though it was, had greater trouble. A paired t-test showed a significant effect of segment for both #_ln_# and #_nl_# type words, with subjects more frequently using a nasal sound than a lateral (**Table 8**). This production of only one sound where English native speakers would use two may indicate an assimilation to the manner of articulation of the other segment [cf. Gordon, 1957, pp. 280–82, in which (n) was deleted before other sonorants, including (l) and the presence of (r) with (l) or (n) led to (ll) and (nn)]. Alternatively, SWM speakers' tendency to produce only a single sound in English words may be due to phonotactic constraints from Mandarin syllabic structure which does not allow (l) in a coda nor (l) and (n) to occur next to each other. English words with a double spelled <nn> or <ll>, which required only a single sound, were produced much more accurately, suggesting that double spellings which do not indicate two different sounds may promote more accurate production for /l/ and for /n/.
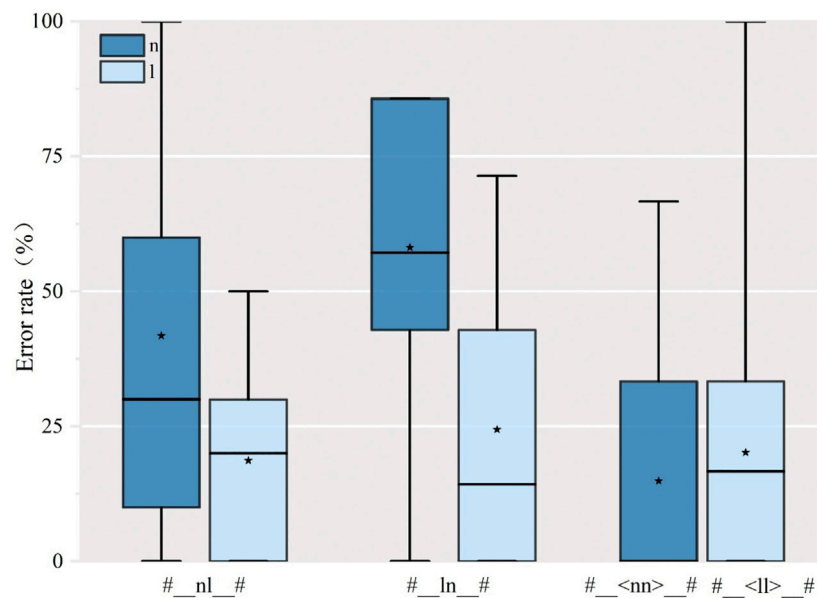
FIGURE 10 | Error rate of production for the English words with the syllable /-nl-/, /-ln-/ and <||>, <nn>.

TABLE 8 | *t*-test results comparing the type of #_nl_# and #_ln_# words.

Comparison	M	SD	t	df	p
#_nl_#→l	0.197	0.169	−2.177	17	0.044
#_nl_#→n	0.441	0.340			
#_ln_#→l	0.230	0.219	−4.408	17	0.000
#_ln_#→n	0.612	0.207			

Notes: #_nl_#→l means the type of #_nl_# words only produce a single lateral sound.
#_nl_#→n means the type of #_nl_# words only produce a single nasal sound.
Same pattern with the type of #_ln_# words.

DISCUSSION

Research Questions

The first research question looked at the effect of L1 dialect on production of /l/ and /n/ in Standard Mandarin. Descriptions of SWM recognize difference within the variety regarding the phonological status of /l/ and /n/ (Li, 2004; Qian, 2010). Some areas have /n/ but not /l/, others the opposite, and yet others have a phonological contrast between the two sounds. Our findings confirm these descriptions of /l/-/n/ variation in SW Mandarin. All subjects who were from Yunnan province showed no difficulty in producing both (n) and (l). In contrast, the subjects who were from the /n/ dialect area showed more difficulty producing (l) than (n), as should be expected. Similarly, the subjects from an attested /l/ dialect area had the opposite pattern, with (n) showing a higher error rate than (l). This indicates that subjects' L1 phonological system influenced their productions in their additional languages. Those with a single category were more accurate in the pronunciation of items that matched the category in their dialect. However, they were also quite accurate overall, and their mispronunciations of /l/ and /n/ were a small minority of the total productions.

The second research question asked about the effects of environment on /l/ and /n/ production in Standard Mandarin. Environment was defined as differences in tone, whether the words were pronounced with or without a competitor sound in phrases, and the effects of the Sihu rhyming system in Mandarin, whose rhyme differences include four categories: rhymes beginning with (i), with (u), with (y), and with non-high vowels. The findings showed that the tone associated with the production of each word of did not affect (l) and (n) production accuracy. The presence or absence of a competitor sound in a phrase showed greater accuracy with /n/ when there was a competitor, and less accuracy for /l/. Finally, the effect of following vowel created different error patterns for (n) and (l) production. For a rhyme beginning with (i), both sounds had similar error rates [18.05% for (n), 21.45% for (l)], but for the other three rhymes, the error rates for (n) production were considerably higher than for (l). This suggests there may be some speakers for whom environment affects their production, and some studies of SWM have indicated that speakers have a phonological contrast between (n) and (l) before high front vowels but not elsewhere (Qian, 2010). Our data do not allow us to confirm this.

The third research question asked about the phonological environments most likely to elicit /n/ and /l/ pronunciation confusions in English. There was no clear effect of environment, but because our subjects included speakers from two phonological systems in uneven numbers, we can only offer descriptive data for the full group. The three initial /l/ environments had error rates near 30% or higher, while all of the /n/ environments had error rates at half the frequency of the pronunciation of /l/. This indicates that environment did not clearly affect the accuracy of the pronunciation of either sound in English. This is perhaps not surprising since our competitor

sounds were in coda position and often in unstressed syllables. Mandarin varieties allow /n/ in codas [though their realization varies between (+consonantal) and (−consonantal)] but do not allow /l/ in coda position (Zhang and Levis, 2020). Child language acquisition shows a tendency to assimilate onset consonants to following consonants (e.g., yellow pronounced as *lellow, as in, for example, Gillespie and Greenberg, 2017). However, studies on tone perception House (1996) have shown that there are different types and amounts of information available from onsets and codas and the two do not consistently affect each other. For future research, initial /n/ and /l/ words could likely include all environments without affecting findings.

It should be pointed out that our SWM subjects correctly produced (n) and (l) most of the time. However, when listening to such speakers, it is our experience that confusion between /l/ and /n/ seems extremely common in spoken production. This may be because both /n/ and /l are frequently-occurring consonants in English (Hayden, 1950), and that listeners notice deviations not correct productions. An analogy to overregularization in child language is instructive. In a corpus of child language, Marcus et al. (1992) found that overregularized forms (e.g., flied for the past tense flew) occurred in only five percent of possible environments. Adult speakers notice these forms in child language and believe they are far more common than they are. (When our linguistics students are asked to estimate the frequency of such forms, giving them four answers, 5, 25, 45 and 65%, they usually pick the two highest frequencies.) Similarly, difficulties in distinguishing /l/ and /n/ in English speech may seem more common than they are. It may be that SWM speakers are using knowledge that /l/ and /n/ are different sounds in other languages to provide imperfect production of both sounds. Archibald (2005) argues that L2 learners can redeploy phonological knowledge from the features of their L1 for more accurate L2 production. For example, Japanese speakers do not have a distinction between /b/ and /v/ but do distinguish stops from continuants (Brown, 2000). As a result, they can use this knowledge of featural distinctions from their L1 to learn an unfamiliar distinction in another language using the same features. The overall accuracy of production suggests that SWM speakers may be able to use some knowledge from their L1 to produce the sounds in their L2 and L3. We cannot know whether SWM speakers similarly make use of their language learning experience (and perhaps the knowledge that their variety is marked in its pronunciation of /l/ and /n/) to pronounce phones more accurately in their L2 and L3. Brown (1998) suggests as much for Japanese speakers when she says that Japanese speakers' performance on the picture task (roughly 60% correct) is, on the whole, better than their performance on the auditory task (roughly 30% correct) (p. 169). If this is the case, it would suggest that in a task in which attention to their production is maximized (as in the reading tasks we used, see Brown, 1998), language learners may be able to demonstrate greater production accuracy than perception accuracy. This question must await perceptual studies.

The fourth research question examined the comparison of SWM speakers' /l/ and /n/ mispronunciations when reading in Standard Mandarin and English. This question sought to establish how the speakers' L1 affected the accuracy of their production in their other languages which had both phonemes. The results showed that the frequency of errors for Standard Mandarin was never more than 15%. Somewhat surprisingly, the error rate increased when the subjects read phrases with only one /l/ or /n/ word, while it decreased when they read phrases with two target sounds. Thus, phrases with competitor sounds were associated with greater accuracy. We expected that such competition would make mispronunciations more common, but this was not supported by the data.

For the English word reading, the /l/ and /n/ mispronunciation rate was higher than for Standard Mandarin, with a significant difference between the accuracy of production of /l/ in English and in Standard Mandarin. This is somewhat surprising. For Standard Mandarin words, both the /l/-only group and /n/-only group (refer to **Figure 2**) demonstrated greater difficulty with the production of (l), but in English, the number of /l/ errors was greater than in Mandarin while the number of /n/ errors remained similar for the L2 and L3 production. This suggests that even for /l/-only subjects, the production of (l) was more difficult.

Finally, we looked at the production accuracy for medial /l/ and /n/, both in words where the sounds were represented by orthographically doubled letters with a single sound and when both sounds were expected to be produced separately in sequence. Words with orthographic doubling of /l/ were produced more accurately than initial /l/ words, and those with medial <nn> were produced with accuracy similar to words with initial (n). In contrast, words with medial <ln> and /<nl> (e.g., walnut, only) were almost never correctly pronounced, with walnut type words being more difficult. In most mispronunciations, subjects produced only one of the sounds, more commonly (n). These words were different from the other words in that their syllable structure included a coda consonant in the first syllable followed by an onset consonant in the second syllable. When the coda was (l) (as in walnut), it involved a violation of the phonotactics of SWM. Zhang and Levis (2020) showed that final /l/ in English is almost never pronounced by SWM speakers, which may be the reason that word internal coda (l) would also be mispronounced. Although coda (n) is licensed in Mandarin, it is often produced as a nasalized vowel, and its production as a consonant occurs only about half the time. In general, Mandarin most commonly is produced with CV syllables, and in this study the two separate consonants in the middle of the word seemed to be assimilated to a single consonant realization in line with Mandarin syllable structure. In our results, this medial consonant is most likely to be (n). Although most of our subjects came from the /l/-only dialect area, and we would expect that they would favor medial (l) rather than (n), this was not the case. More importantly, because these medial combinations of (n) and (l) are almost certain to be mispronounced by SWM speakers, it would be valuable to know whether the mispronunciations affect how English listeners understand the words. The initial /l/-/n/ contrast has a high functional load (Catford, 1987), but we do not have equivalent measure for loss of a segment in the middle of a

word. The earlier anecdote about pronouncing walnuts as wallets creates a different word that caused confusion for the listeners, but not all examples are so clear. Pronouncing only as either (oni) or (oli) while preserving the original stress pattern may be fully intelligible to a listener, and there may be no need in most cases to pronounce both sounds. Instead, it may be better to focus attention on particular words that cause loss of understanding.

/l/-/n/ in Cantonese and in Southwestern Mandarin

There is a substantially larger amount of research on /l/ and /n/ in Cantonese speakers' (l)-(n) production than in SWM, but our findings make clear that the lack of an /l/-/n/ contrast in the two Chinese languages follow different patterns. Cantonese, especially Hong Kong Cantonese, is in the process of a merger between the two sounds that suggests a complementary distribution, with (l) occurring in onset and (n) in coda position (Ng, 2017). This merger is more evident in younger speakers (Yeung, 1980; Tong and James, 1994), but HKC speakers perceive the differences between the two sounds, indicating that this advanced merger remains incomplete (Chan, 2011). Additionally, Cantonese speakers may infrequently pronounce (l) as (ɹ) (Chan, 2010), an error we found even more rarely in our SWM data. In SWM, any merger of the two sounds (l) and (n) took place in more distant history; some varieties of SWM currently have /l/ and others /n/, but the /l/ was a more difficult sound in English for all SWM speakers. There was also difficulty in our agreeing on identification of some English productions because the subjects sometimes appeared to produce a co-articulated sound with both nasal and lateral elements, sometimes leading to confusion for us as researchers.

Indeed, some productions defied easy categorization in the English data. One example is the word "nock" produced by subject D15 (provided in the **Supplementary Material** on the journal site). One author was confident that the initial sound was a lateral, but the other was certain it was a nasal. The acoustic measurements could not completely clarify our disagreement as the productions had acoustic evidence of both phonetic features. This indicates the need for an analysis of SWM (l) and (n) in terms of phonological features, especially because both /n/ and /l/ are ambiguous in regards to the feature (continuant), and their specific feature settings may be different in different languages (Mielke, 2005). It may also be the case that the features (nasal) and (lateral) are not distinguished for SWM speakers, causing us as listeners to hear their productions in terms of our own phonemic systems, which may not be equivalent in how the features distinguishing /n/ and /l/ are realized. As mentioned earlier, some researchers have argued that hybrid productions that include aspects of both lateral and nasal sounds occur within some Mandarin varieties (Chan, 1987; Soejima et al., 1990). How this would be described in terms of features is not clear, but it is likely that such productions would include both (nasal) and (lateral) features. This intriguing possibility must, however, be left for future research.

Influences on L2 and L3 Pronunciation

Although we do not have evidence of L1 patterns of /l/-/n/ productions for SW Mandarin itself, there is suggestive evidence that the L1 system affects production in both the L2 (Standard Mandarin) and L3 (English). The first piece of evidence is that the effect of the sub-variety was similar for both L2 and L3. Speakers with only /n/ in their L1 had higher numbers of /l/ errors in both the L2 and L3, and those with only /l/ in their phonological inventory had more /n/ errors in L2 and L3. This indicates a consistent effect of the L1 on both additional languages. A second piece of evidence is that /l/ appeared to be more challenging than /n/ for both L2 and L3 production, suggesting the same influence on both additional languages. Despite differences in frequency of errors, there was a strong correlation between the performance in both the L2 and L3. Those subjects with better accuracy in Standard Mandarin were likely to have better performance in English, and those with worse performance in English were likewise less accurate in Standard Mandarin. Evidence against the influence of the L2 on L3 can be seen in how English error rates for /l/ were higher than the frequency of Standard Mandarin errors. If L2 production had been influential on L3 production, we would expect fewer English errors than we in fact found. Instead, the language distance (Llama et al., 2010) between SWM and Standard Mandarin, two historically related varieties, was associated with greater accuracy in Standard Mandarin production than in English, which is more typologically distant.

The differing accuracy rates may also occur because of proficiency differences and different phonotactic constraints in the L2 and L3. SWM speakers learn Standard Mandarin from an earlier age and the language is presented consistently in the school setting. English is introduced later and is used primarily within English classes. SWM and Standard Mandarin share the same syllable structures, the same possible environments for both (n) (onset and coda) and (l) (onset only), and little difference in basic vocabulary, making the words produced more familiar. In contrast, the production of (n) and (l) in English occurs in unfamiliar linguistic environments (initial, in clusters, medially, and final) especially when the sounds were produced in clusters. The syllable structure and phonotactics of English are more complex than those of Mandarin, and this may make the production of (l) and (n) more challenging overall because /l/ and /n/ get pronounced in more environments and include different allophones.

Finally, SWM learners of English also have a higher learning burden for English, not only phonologically but also lexically and orthographically, contending with both more complex phonotactics and a more challenging knowledge of vocabulary. This suggests that when speaking English, SWM speakers may pay less attention to particular segments because of the extra attention needed to attend to meaning. In addition, English words are represented with a different writing system, which may add to the burden of accurate pronunciation. Latin orthography in the form of pinyin is used for elements of Mandarin reading, but its orthographic system does not have the same sound-spelling correspondence used in English. Even though <l> and <n>

often reflect the corresponding phonemes, English's notably opaque orthography adds to the burden of pronouncing words that learners may or may not have ever heard before.

Implications for Pronunciation Teaching

The /l/-/n/ contrast, as mentioned earlier, is similar in functional load to contrasts such as /l/-/ɹ/ and /p/-/b/, which are at the highest level of functional load in English (Catford, 1987; Brown, 1988). This means that it is very likely that confusions of these two phonemes in English will lead to misunderstandings, to challenges in processing speech, and to increased perceptions of accentedness (Munro and Derwing, 2006). Higher functional load is well-established as an important criterion for comprehensibility and accentedness in English (Munro and Derwing, 2006; Suzukida and Saito, 2019), but it is also likely that this contrast is equally important for learning other major languages such as French, Swedish, Russian, and Spanish, all of which have the /l/-/n/ contrast. Thus this challenge for SWM speakers (and Cantonese speakers) in speaking English may also be relevant for a variety of other additional languages.

Our findings indicate that /l/ may be a more challenging sound for SWM speakers in English as /l/ words had consistently higher error rates than /n/ words, and in words with both sounds (e.g, only, walnut), subjects were more likely to delete (l) and pronounce the (n). Nonetheless, the degree of difference between /n/ and /l/ accuracy for different subjects may not be relevant from a teacher's point of view since both the /n/ group and /l/ group had a similar combined accuracy on /n/ and /l/ words together. Although they had different proportions of /l/ and /n/ mispronunciation, from the viewpoint of a teacher, they would have similar numbers of errors. In English, /n/ and /l/ are two of the five most common consonants (Hayden, 1950), which means that listeners will be regularly confronted by word identification decisions when speaking with a SWM speaker who does not consistently produce a difference between the sounds. For SWM speakers, this may indicate a need for not only production but also perception training. If the difficulties that SWM speakers have in production are related to perception, then robust perceptual training using High Variability Phonetic Training (HVPT) may be required to help develop or strengthen distinct phonetic categories for /l/ and /n/. Although there is evidence that Cantonese speakers perceive the differences between (l) and (n) (Chan, 2011), because of the phonological distribution of the two sounds, there is no reason to assume this is the case for SWM speakers. Qian (2018) looked at the /l/-/n/ contrast in her HVPT study of Chinese learners. For her subjects, which included some SWM speakers, HVPT training appeared to help the few speakers who had this challenge. A study involving only SWM speakers would be valuable in determining the extent to which SWM speakers can improve their perception of the two sounds.

The results suggest that there are easier and more difficult environments for teaching. Single medial /l/ and /n/ seemed to be the most successfully pronounced and may be a good place to start with production and perception practice, while medial <ln> and <nl> were almost always mispronounced. These frequent mispronunciations of medial sound pairs may suggest that these

should be a priority, but an intelligibility-based approach to teaching (Levis, 2018) asks for evidence that errors affect understanding. However, we have no evidence that producing only one sound where two are expected affects understanding in the way that functional load predicts problems for initial /l/ and /n/, which are likely to be more important because initial mispronunciations can lead listeners to access the wrong cohort of potential words (Marslen-Wilson and Welsh, 1978; Zielinski, 2008).

Limitations

A clear limitation of our study is that we had uneven numbers of subjects from different sub-varieties. It would be helpful to have larger and more equal numbers of subjects controlled for sub-variety. The exploratory nature of our study and the uneven numbers only made it possible for us to suggest trends in the production of /l/ and /n/. Another limitation was in the nature of the data we collected. There was no free speech, which may have provided different frequencies for confused sounds. This could also make it possible to quantify the comprehensibility of SWM speech by allowing listeners to rate the speech. Third, in collecting our recordings, we did not include distractor items though it is not clear that this affected how our subjects attended to the task. Even though all the words included (l) and (n), several subjects still asked the first author about the purpose of the study after reading the items. In addition, the uneven numbers of items produced added complications to our ability to compare frequencies. Some had 34 items (initial /n/ and /l/ for English) while other had as few as seven items or as many as 400 (Standard Mandarin words). Larger numbers of words almost certainly guaranteed greater numbers of unknown items, perhaps causing unforeseen difficulties in lexical access. Finally, our study did not include perception data. Because perception data is a better measure of whether there are intractable limitations in the L2 phonological system (Brown, 1998; Brown, 2000), and production data may not be directly related to perception, perception data would have allowed more confident interpretations of the production data.

Future Research

As mentioned above, it is very important that there be perception studies about how well SWM speakers hear (l) and (n). We know that the perception of even very difficult L2 contrasts can be improved (e.g., Bradlow et al., 1999), and that improved perception can also lead to improved production (Sakai and Moorman, 2018), although this is not always the case (Kartushina et al., 2015). If /l/-/n/ difficulties in production have a perceptual component, then high variability phonetic training approaches can be employed to build new category boundaries (Barriuso and Hayes-Harb, 2018) and complement production practice.

Another type of study that would be valuable would be an acoustic analysis of /l/ and /n/ production by SWM speakers. There are suggestions in previous research (Soejima et al., 1990) that SWM speakers may produce a sound that has both nasal and lateral realizations (e.g., a lateral nasal), indicating that a better analysis of the /l/-/n/ merger would be in terms of features

rather than segments, especially for items that seemed to have features associated with phonologically distinct segments in languages like English. This would require a careful acoustic analysis that looked at productions that were in-between the category boundaries of the researchers. We identified quite a few potential tokens in our own listening, and they invariably led us as English-L1 and Standard Mandarin-L1 listeners to disagree on their categorization. If SWM speakers indeed have a sound that has characteristics of both nasals and laterals, it would suggest that /l/-/n/ problems are more like that attested for Japanese speakers with /l/-/ɹ/. Alternatively, we could be seeing pronunciations that result from speakers with one category speaking a language with a new category. Additionally, in regard to only and walnut type words, It may be that the subjects produced the singleton sound with greater duration, that is, as a geminate, to acoustically signal a difference between words with a single sound and these words. This would require production data for words like whining and whaling, in which there was only one medial consonant letter.

It would also be interesting to examine the L2 speech of different age groups. The learning of Standard Mandarin has become more standardized in China's education system over many decades, and our subjects began learning Standard Mandarin at the beginning of their schooling. Subjects who are older (e.g., over 40 years old) are less likely to have received early training in Standard Mandarin or English, and the effects of their L1 may be greater.

CONCLUSION

The SWM lack of contrast between /l/ and /n/ is a problem that helps explore the interface between L2 phonology and L2 pronunciation and is important for a number of reasons. First, from a point of view of L2 phonology, this study addresses a phonemic contrast that is common in a wide variety of languages (Yip, 2011) and is thus relevant to a variety of L2 learning contexts. SWM is unusual in not having such a distinction. Since our production data do not make clear the extent to which (l) and (n) represent two allophones of one phoneme in SWM, or whether they represent a sound in SWM sub-varieties that is neither /l/ nor /n/ (Soejima et al., 1990), data that bears on this question would be valuable for studying the phonetics and phonology of SW Mandarin.

From an L2 pronunciation point of view, the sounds in this study are important to L2 learners from a large population. More than 85,000,000 people speak Cantonese around the world, most of them L1 speakers in China and Hong Kong (Ethnologue, 2021), and up to 270,000,000 speakers speak some variety of SW Mandarin (Chinese Academy of Social Sciences, 2012). The production of these sounds in English thus becomes important because English is a required subject throughout their education beginning at age nine.

Our starting point for this study came from our interests as language teachers, which led us to examine how SWM speakers produced words with /l/ and /n/ in their additional languages. We were especially interested in the frequency and distribution of the subjects' /l/-/n/ errors in English because of the likelihood of such errors leading to unintelligibility in SWM speakers' English. It is our hope that this study provides a beginning that will lead to a more accurate understanding of where and how often such errors occur, and that our results may inform English pronunciation teachers' work with SWM learners.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Qufu Normal University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

WZ collected the recordings for the study, did a large part of the data analysis, and was equally involved in writing the paper. JL contributed to the data analysis, and was equally involved in writing the paper.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm.2021.639390/full#supplementary-material

REFERENCES

Amaro, J. C. (2017). Testing the Phonological Permeability Hypothesis: L3 Phonological Effects on L1 versus L2 Systems. *Int. J. Bilingualism* 21 (6), 698–717. doi:10.1177/1367006916637287

Ao, R., and Low, E. L. (2012). Exploring Pronunciation Features of Yunnan English. *English Today* 28 (3), 27–33. doi:10.1017/s0266078412000284

Aoyama, K., Guion, S. G., Flege, J. E., Yamada, T., and Akahane-Yamada, R. (2008). The First Years in an L2-Speaking Environment: A Comparison of Japanese Children and Adults Learning American English. *IRAL - Int. Rev. Appl. Linguistics Lang. Teach.* 46 (1), 61–90. doi:10.1515/iral.2008.003

Archibald, J. (2005). Second Language Phonology as Redeployment of L1 Phonological Knowledge. *The Can. J. Linguistics / La revue canadienne de linguistique* 50 (1), 285–314. doi:10.1353/cjl.2007.0000

Au, Y. N. A. (2011). The Markedness Theories and the Relationship between /n/ and /l/ in the English Syllable of Cantonese Speakers. Proceedings of the 16th Conference of Pan-Pacific Association of Applied Linguistics, August 2011, pp. 140–141.

Baker, H. D. R. (1993). Language atlas of China. Gen Ed. (Australia), S. A. Wurm, B. T'sou, D. Bradley; (China) Li Rong, Xiong Zhenghui, Zhang Zhenxing. [In two parts, boxed]. (Pacific Linguistics, Series C, no. 102). 36 maps, 22 sheets text material + addenda et corrigenda. Canberra: Australian Academy of the Humanities and the Chinese Academy of Social Sciences [and] Department of Linguistics, Research School of Pacific Studies, ANU; Hong Kong: Longman Group (Far East), 1987, 1989. *Bull. Sch. Oriental Afr. Stud.* 56 (2), 398–399. doi:10.1017/S0041977X0000598X

Bao, H. Q., and Lin, M. C. (2014). *ShiYan YuYingXue GaiYao (Revised Edition)*. Beijing: Peking University Press.

Barriuso, T. A., and Hayes-Harb, R. (2018). High Variability Phonetic Training as a Bridge from Research to Practice. *CATESOL J.* 30 (1), 177–194.

Bent, T., Bradlow, A. R., and Smith, B. (2007). "Phonemic Errors in Different Word Positions and Their Effects on Intelligibility of Non-native Speech," in *Second-language Speech Learning: The Role of Language Experience in Speech Perception and Production: A Festschrift in Honour of James E. Flege*. Editors O.-S. Bohn and M. J. Munro (Amsterdam: John Benjamins), 331–347. doi:10.1075/lllt.17.28ben

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., and Tohkura, Y. i. (1999). Training Japanese Listeners to Identify English /r/and /l/: Long-Term Retention of Learning in Perception and Production. *Perception & Psychophysics* 61 (5), 977–985. doi:10.3758/bf03206911

Brown, A. (1988). Functional Load and the Teaching of Pronunciation. *TESOL Q.* 22 (4), 593–606. doi:10.2307/3587258

Brown, C. A. (1998). The Role of the L1 Grammar in the L2 Acquisition of Segmental Structure. *Second Lang. Res.* 14, 136–193. doi:10.1191/026765898669508401

Brown, C. (2000). "The Interrelation between Speech Perception and Phonological Acquisition from Infant to Adult," in *Second Language Acquisition and Linguistic Theory*. Editor J. Archibald (Malden, MA: Blackwell), 4–63.

Cabrelli Amaro, J. (2012). "L3 Phonology," in *Third Language Acquisition in Adulthood*. Editors J. Cabrelli Amaro, S. Flynn, and J. Rothman (Amsterdam/Philadelphia: John Benjamins), 33–60. doi:10.1075/sibil.46.05ama

Catford, J. C. (1988). "Functional Load and Diachronic Phonology," in *The Prague School and its Legacy*. Editor Y. Tobin (Amsterdam: John Benjamins), 3–19. doi:10.1075/llsee.27.04cat

Catford, J. C. (1987). "Phonetics and the Teaching of Pronunciation: a Systemic Description of English Phonology," in *Current Perspectives on Pronunciation: Practices Anchored in Theory*. Editor J. Morley (Washington, DC: TESOL Press), 87–100.

Chan, A. Y. W. (2010). Advanced Cantonese ESL Learners' Production of English Speech Sounds: Problems and Strategies. *System* 38 (2), 316–328. doi:10.1016/j.system.2009.11.008

Chan, A. Y. W., and Li, D. C. S. (2000). English and Cantonese Phonology in Contrast: Explaining Cantonese ESL Learners' English Pronunciation Problems. *Lang. Cult. Curriculum* 13 (1), 67–85. doi:10.1080/07908310008666590

Chan, A. Y. W. (2014). The Perception and Production of English Speech Sounds by Cantonese ESL Learners in Hong Kong. *Linguistics* 52 (1), 35 – 72. doi:10.1515/ling-2013-0056

Chan, A. Y. W. (2011). The Perception of English Speech Sounds by Cantonese ESL Learners in Hong Kong. *TESOL Q.* 45 (4), 718–748. doi:10.5054/tq.2011.268056

Chan, M. K. M. (1987). Post-stopped Nasals in Chinese: An Areal Study. *UCLA Working Pap. Phonetics* 68, 73–120.

Chen, X. (2013). *Study on the Origin and Change of the Phonetic System of Yunnan Dialect*. Dissertation. Tianjin, China: Nankai University.

Chinese Academy of Social Sciences (2012). *Zhōngguó Yǔyán Dìtú Jí (Dì 2 Bǎn): Hànyǔ Fāngyán Juàn [Language Atlas of China. in Chinese Dialect Volume]*. 2nd edition (Beijing: The Commercial Press).

Deterding, D. (2010). ELF-based Pronunciation Teaching in China. *Chin. J. Appl. Linguistics* 33 (6), 3–15.

Deterding, D. (2006). The Pronunciation of English by Speakers from China. *Eww* 27 (2), 175–198. doi:10.1075/eww.27.2.04det

Deterding, D., Wong, J., and Kirkpatrick, A. (2008). The Pronunciation of Hong Kong English. *Eww* 29 (2), 148–175. doi:10.1075/eww.29.2.03det

Duanmu, San. (2000). *The Phonology of Standard Chinese*. New York, Oxford: Oxford University Press.

Ethnologue (2021). Chinese, Yue. Availableat: https://www-ethnologue-com.eu1.proxy.openathens.net/language/yue (Accessed June 1, 2021).

Feng, S. L. (2016). Mandarin Chinese Is a Stress Language. *YUYAN KEXUE* (05), 449–473.

Gillespie, L. G., and Greenberg, J. D. (2017). Empowering Infants' and Toddlers' Learning through Scaffolding. *YC Young Child.* 72 (2), 90–93.

Gordon, E. V. (1957). *An Introduction to Old Norse*. Oxford, UK: Oxford University.

Goto, H. (1971). Auditory Perception by normal Japanese Adults of the Sounds "L" and "R". *Neuropsychologia* 9, 317–323. doi:10.1016/0028-3932(71)90027-3

Hammarberg, B. (2001). "Chapter 2. Roles of L1 and L2 in L3 Production and Acquisition," in *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives*. Editors J. Cenoz, B. Hufeisen, and U. Jessner (Tonawanda, NY, United States: Multilingual Matters Ltd), 21–41. doi:10.21832/9781853595509-003

Han, S. X. (2014). A Research into the Negative Transfer of Gansu Dialect on the English Phonetic Learning. *J. Longdong Univ.* 25 (6), 45–47.

Hayden, R. E. (1950). The Relative Frequency of Phonemes in General-American English. *WORD* 6 (3), 217–223. doi:10.1080/00437956.1950.11659381

Hazan, V., Sennema, A., Iba, M., and Faulkner, A. (2005). Effect of Audiovisual Perceptual Training on the Perception and Production of Consonants by Japanese Learners of English. *Speech Commun.* 47 (3), 360–378. doi:10.1016/j.specom.2005.04.007

House, D. (1996). Differential Perception of Tonal Contours through the Syllable. Proceedings of Fourth International Conference on Spoken Language Processin. *ICSLP*, 96. IEEE, 2048–2051.

Hu, F. (2007). *International Congress of Phonetic Sciences XVI*, 1405–1408. Availableat: http://www.icphs2007.de/conference/Papers/1127/1127.pdf .Post-oralized Nasal Consonants in Chinese Dialects—Aerodynamic and Acoustic Data

Huang, B. R., and Liao, X. D. (2011). *Modern Chinese*. Beijing, China: Higher Education Press.

Hung, T. T. N. (2000). Towards a Phonology of Hong Kong English. *World Englishes* 19 (3), 337–356. doi:10.1111/1467-971x.00183

IIE (2019). The Power of International Education: Number of International Students in the United States Hits All-Time High. Availableat: https://www.iie.org/Why-IIE/Announcements/2019/11/Number-of-International-Students-in-the-United-States-Hits-All-Time-High (Accessed March 2, 2020).

Kang, O., and Moran, M. (2014). Functional Loads of Pronunciation Features in Nonnative Speakers' Oral Assessment. *Tesol Q.* 48 (1), 176–187. doi:10.1002/tesq.152

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., and Golestani, N. (2015). The Effect of Phonetic Production Training with Visual Feedback on the Perception and Production of Foreign Speech Sounds. *The J. Acoust. Soc. America* 138 (2), 817–832. doi:10.1121/1.4926561

Koffi, E. (2019). An Acoustic Phonetic Account of the Confusion between [l] and [n] by Some Chinese Speakers. *Linguistic Portfolios* 8 (6), 64–73.

Kurpaska, M. (2010). *Chinese Language(s): A Look through the Prism of the Great Dictionary of Modern Chinese Dialects*. Berlin, New York: De Gruyter Mouton.

Ladefoged, P., and Johnson, K. (2011). *A Course in Phonetics*. 6th Ed.. Boston: Wadsworth.

Ladefoged, P., and Maddieson, I. (1996). *The Sounds of the World's Languages*. MA: Blackwell: Cambridge.

Lan, Q. Y. (1995). Hechi Dialect Phonology. *Acad. Forum* 5, 51–56.

Leung, M., and Sharma, Y. (2020). International Student Flows Extend Fallout from Coronavirus Outbreak. Availableat: https://www.universityworldnews.com/post.php?story=20200129174911175 (Accessed to March 2, 2020).

Levis, J. (2018). *Intelligibility, Oral Communication, and the Teaching of Pronunciation*. New York: Cambridge University Press.

Li, B., and Hua, C. C. (2014). Effects of Visual Cues on Perception of Non-native Consonant Contrasts by Chinese EFL Learners. *Asian EFL J.* 16 (1), 271–295.

Li, L. (2009). Classification/distribution of Southwest Mandarin. *Dialects* 1, 72–87.

Li, M. (2017). *Study on the Phonetic System of the Dialects in the Area of Nanchong, Sichuan Province*. Dissertation. Chengdu, China: Sichuan Normal University.

Li, Q. Q. (2002). *Jishou Fāng Yán Yán Jiū [Study of Jishou Dialect]*. Beijing: The Ethnic Publishing House.

Li, X. (2004). *Comprehensive Phonetic Study of Southwestern Mandarin*. Dissertation. Shanghai, China: Shanghai Normal University.

Li, Y. M., Zhao, S. J., and He, L. (2020). The Practice of and Reflections on "Epidemic Language Service Corps". *Chin. J. Lang. Pol. Plann.* 27 (3), 23–30.

Lin, T., Wang, L. J., and Wang, Y. J. (2013). *YuYinXue JiaoCheng (Revised Edition)*. Beijing: Peking University Press.

Llama, R., Cardoso, W., and Collins, L. (2010). The Influence of Language Distance and Language Status on the Acquisition of L3 Phonology. *Int. J. Multilingualism* 7 (1), 39–57. doi:10.1080/14790710902972255

Luo, Y. H. (2016). Phonology of Guizhou Zunyi Dialect. *Mod. Chin.* 05, 43–45.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., Xu, F., et al. (1992). Overregularization in Language Acquisition. *Monogr. Soc. Res. Child Development* 57, i–178. doi:10.2307/1166115

Marslen-Wilson, W. D., and Welsh, A. (1978). Processing Interactions and Lexical Access during Word Recognition in Continuous Speech. *Cogn. Psychol.* 10 (1), 29–63. doi:10.1016/0010-0285(78)90018-x

Mielke, J. (2005). Ambivalence and Ambiguity in Laterals and Nasals. *Phonology* 22 (2), 169–203. doi:10.1017/s0952675705000539

Ming, S. R. (2005). A Study on the Law of Correspondence between the Initial Consonants in Bijie Dialect and the Initial Consonants of Beijing Pronunciation. *J. Guizhou Education Inst. (Social Science)* 01, 75–77+103.

Ming, S. R. (1997). The Vowels, Rhymes and Tones of the Bijie Dialect. *J. Guizhou Education Univ.* 02, 82–84.

Modern Chinese Teaching and Research Program, Chinese Department, Peking University (2002). *Modern Chinese*. Beijing: The Commercial Press.

Munro, M., Derwing, T., and Thomson, R. (2015). Setting Segmental Priorities for English Learners: Evidence from a Longitudinal Study. *Int. Rev. Appl. Linguistics Lang. Teach.* 53 (1), 39–60. doi:10.1515/iral-2015-0002

Munro, M. J., and Derwing, T. M. (2006). The Functional Load Principle in ESL Pronunciation Instruction: An Exploratory Study. *System* 34 (4), 520–531. doi:10.1016/j.system.2006.09.004

Ng, C. L. C. (2017). Merger of the Syllable-Initial [n-] and [l-] in Hong Kong Cantonese. (Outstanding Academic Papers by Students (OAPS), City University of Hong Kong). Availableat: http://lbms03.cityu.edu.hk/oaps/lt2017-4235-ncl190.pdf.

Pater, J., and Werle, A. (2003). Direction of Assimilation in Child Consonant harmony. *The Can. J. Linguistics/La revue canadienne de linguistique* 48 (2), 385–408. doi:10.1353/cjl.2004.0033

Pennington, R., and Saunders, J. (2013). *Introduction to the Documentary Corpus of Guiliu, a Southwestern Mandarin Dialect*. Course paper.

Qian, M. (2018). *An Adaptive Computational System for Automated, Learner-Customized Segmental Perception Training in Words and Sentences: Design, Implementation, Assessment*. Doctoral Dissertation. Ames, IA, United States: Iowa State University

Qian, Z. Y. (2010). *A Study of Mandarin Chinese Dialect*. Jinan, China: Qilu Publishing House.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. Computer Program. Version 2.9.2. Vienna, Austria: R Foundation for Statistical Computing. Availableat: http://www.R-project.org.

Raver-Lampman, G., Toreno, F., and Bing, J. (2015). An Acceptable Non-alveolar Japanese /s/. *Jslp* 1 (2), 238–253. doi:10.1075/jslp.1.2.05rav

Richards, M. (2012). "Helping Chinese Learners Distinguish English /l/ and /n/," in *Proceedings Of the 3rd Pronunciation In Second Language Learning And Teaching Conference, Sept. 2011*. Editors J. Levis and K. LeVelle (Ames, IA: Iowa State University), 161–167.

Sakai, M., and Moorman, C. (2018). Can Perception Training Improve the Production of Second Language Phonemes? A Meta-Analytic Review of 25 Years of Perception Training Research. *Appl. Psycholinguistics* 39 (1), 187–224. doi:10.1017/s0142716417000418

Sewell, A. (2017). Functional Load Revisited. *Jslp* 3 (1), 57–79. doi:10.1075/jslp.3.1.03sew

Soejima, K., Imagana, H., and Kiritani, S. (1990). Alternation between Stop Nasal and (Nasalized) Flap or Lateral. *Ann. Bull. RILP* 24, 131–144.

Su, P. (2010). Photosynthesis of C4 Desert Plants. *Anhui Lit.* 8, 243–259. doi:10.1007/978-3-642-02550-1_12

Sun, Y. (2011). *Studies on Dialects of Southwestern Mandarin in Sichuan Province - Their Phonological Systems and Historical Development*. Dissertation. Hangzhou, China: Zhejiang University.

Suzukida, Y., and Saito, K. (2019). Which Segmental Features Matter for Successful L2 Comprehensibility? Revisiting and Generalizing the Pedagogical Value of the Functional Load Principle. *Lang. Teach. Res.* 136216881985824.

Tong, K. S., and James, G. (1994). *Colloquial Cantonese*. New York: Psychology Press.

Třísková, H. (2011). Chapter Four. The CVX Theory of Syllable Structure. *Archiv orientální* 79 (1), 99–127. doi:10.1163/ej.9789004187405.i-464.24

Wang, H., Wu, D., Chen, H., and Chen, J. C. (2009). Research on Yibin Dialect and Mandarin Chinese in Initial Part and Final Part of a Syllable. *J. Chongqing Univ. Arts Sci. (Social Sci. Edition)* 1, 86–88.

Wang, P. (1981). The Phonological System of Guiyang Dialect. *Dialect* 02, 122–130.

Wang, S. J. (1986). The Phonetic Differences between Qujing Dialect and Mandarin Chinese. *J. Qujing Teach. Coll.* 00, 75–78.

Wei, J. (2018). A Comparative Analysis of Consonants between Mengzi Dialect and Standard Mandarin. *Data Cult. Education* 15, 24–26.

Wong, M. (2008). Can Consciousness-Raising and Imitation Improve Pronunciation?. *Int. J. Learn. Annu. Rev.* 15 (6), 43–48. doi:10.18848/1447-9494/cgp/v15i06/45808

Wrembel, M. (2013). "Foreign Accent Ratings in Third Language Acquisition: The Case of L3 French," in *Teaching and Researching English Accents in Native and Non-native Speakers*. Editors E. E. Waniek-Klimczak and L. Shockey (Berlin: Springer), 31–47. doi:10.1007/978-3-642-24019-5_3

Xiao, Y. F. (1996). Anita Hill Resigns Her Professorship at the University of Oklahoma College of Law. *J. Blacks Higher Education* 01, 69–81. doi:10.2307/2962837

Xin, M. (2013). The Phonetic System of Guizhou Xingyi Dialect. *J. Xingyi Normal Univ. Nationalities* 01, 57–61.

Yeung, S. W. (1980). *Some Aspects of Phonological Variations in the Cantonese Spoken in Hong Kong*. Dissertation. [Hong Kong]: University of Hong Kong.

Yip, M. (2011). "Lateral Consonants," in *The Blackwell Companion to Phonology*. Editors M. van Oostendorp, C. Ewen, E. Hume, and K. Rice (Wiley Blackwell), 1–26. doi:10.1002/9781444335262.wbctp0031

Yuan, B. L. (1996). Anshun Urban Dialect Phonological System. *Anshun Shizhuan Xuebao* 01, 61–73.

Zee, E. (1999). *Chinese (Hong Kong Cantonese). The Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press, 58–60.

Zeng, X. G. (2009). *Nanchong Fāng Yán Yán Jiū [Study of Nanchong Dialect]*. Chengdu: Sichuan People's Publishing House.

Zhang, J. S. (2021). Also On Chinese Word Stress. *Chin. Lang.* 01, 43–55+127.

Zhang, W. (2007). Alternation of [n] and [l] in Sichuan Dialect, Standard Mandarin and English: a Single-Case Study. *Leeds Working in Linguistics and Phonetics* 12, 156–173.

Zhang, W., and Levis, J. (2020). "Southwestern Mandarin Speakers' Production of English Word-Final /l/ and /n/," in *Proceedings Of the 11th Pronunciation In Second Language Learning And Teaching Conference, ISSN 2380-9566, Northern Arizona University, September 2019*. Editors O. Kang, S. Staples, K. Yaw, and K. Hirschi (Ames, IA: Iowa State University), 292–303.

Zhang, Z. R. (2012). Analysis of the Influence of Gansu Dialect on English Pronunciation. *J. Gansu Lianhe Univ. (Social Sciences)* 28 (2), 73–75.

Zheng, Q. J. (1999). *Changde Fāng Yán Yán Jiū [Study of Changde Dialect]*. Zhangsha: Hunan Education Publishing House.

Zhou, R. (2018). A Review on Mandarin Word Stress Study in 60 Years. *Lang. Teach. Linguistic Stud.* (06), 102–112.

Zhou, Y. Y. (2014). *The Experimental Analysis of Dialectal Phonetic Systems and Research on Their Linguistic Geography in Mianyan Sichun*. Dissertation. Chengdu, China: Sichuan Normal University.

Zielinski, B. (2006). The Intelligibility Cocktail: An Interaction between Speaker and Listener Ingredients. *Prospect: Aust. J. TESOL* 21 (1), 22–45.

Zielinski, B. W. (2008). The Listener: No Longer the Silent Partner in Reduced Intelligibility. *System* 36 (1), 69–84. doi:10.1016/j.system.2007.11.004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer, AS, declared a past co-authorship with one of the authors, JL.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.