To appears as: Pelzl, E. (2021) Foreign-accent in second language Mandarin Chinese. In C. Yang (Ed.) *The acquisition of Chinese as a second language pronunciation: Segments and prosody.* Springer.

Foreign accent in second language Mandarin Chinese

Eric Pelzl
The Pennsylvania State University
University Park, Pennsylvania, United States

Abstract

This chapter discusses second language pronunciation of Mandarin from the perspective of the native Mandarin speakers who listen to it. For such listeners, second language Mandarin often bears a noticeable foreign accent. I will provide a framework for defining foreign accent and for distinguishing accented pronunciation from pronunciation errors. I will then review the results of research related to foreign-accented Mandarin and how it affects listeners' judgments, comprehension, and the efficiency with which they process second language Mandarin speech. Naturally, lexical tones will receive special attention in this discussion.

Keywords: Mandarin, second language pronunciation, foreign accent, pronunciation error, tones

1. Introduction

Mandarin Chinese (*Pǔtōnghuà*) speakers often use the phrase *yáng qiāng yáng diào* (洋腔洋调) to describe the speech produced by second language (L2) Mandarin speakers. Ignoring, for the moment, that this phrase may come with some social baggage, its existence shows that native Chinese listeners hear something different in L2—or foreign-accented—speech. Even though listeners are familiar with Cantonese, Shanghainese, Taiwanese and other native Mandarin accents, in some perceptible way foreign-accented speech is different. I know what you're thinking—*it's the tones!* That may be correct, but in this chapter, we aren't going to rush to any conclusions. We will take our time considering the many ways that foreign-accent might be apparent in L2 Mandarin speech and how this impacts listeners.

We will start by defining some of the important qualities of L2 pronunciation, but overall our focus will be on how foreign-accented speech affects native Chinese interlocutors—the common conversation partners of L2 speakers. By approaching L2 pronunciation from the perspective of listeners, we can gain insight into which aspects of pronunciation ought to be prioritized in learning and teaching. In this chapter, I won't attempt to explain *why* L2 accent happens, but interested readers can refer to key theoretical studies considering L2 pronunciation (e.g., Best and Tyler 2007; Escudero and Boersma 2004; Flege 1995; Major 2001).

As a review of research, this one comes with a big caveat—there is not much to review that is specific to L2 Mandarin. Research on accented speech perception and comprehension is only just beginning and, by my count there, are only five existing studies that have directly addressed foreign-accented Mandarin (not including other studies that might appear in this volume). I will review them all in some detail, focusing on ways that we might build on them for future research, but I will also draw heavily on other lines of research on native (L1) and L2 Mandarin speech. I will also draw connections to the much more extensive work that has examined foreign-accented speech in other languages, especially English.

One last note before we get going in earnest. Describing the people who speak with accents in this research is not always straightforward. In places where Mandarin is recognized as an official language, there are many who identify as minority language speakers. For them Mandarin is also a second language that they may only ever master imperfectly. Additionally, given the diversity of Chinese regional languages (Norman 1988), many who identify as L1 Mandarin speakers, also produce the language with 'non-standard' accents and, technically, might be called L2 speakers. For the purposes of this chapter, the L2 speakers we will be thinking about are primarily those who have lived most of their lives outside of Chinese language communities and have learned Mandarin largely as adults. It is this type of learner that we know is very likely to speak with a noticeable *foreign* accent (Flege, Munro, and MacKay 1995). Differences among L2 speakers' native language backgrounds will certainly lead to different qualities of foreign-accent. However, the few currently existing studies on foreign-accented Mandarin include L2 participants from a mix of L1 backgrounds, so we will not narrow in on any specific L1 in this review.

We begin our discussion with an attempt to more clearly define foreign accent.

2 What is foreign-accented Mandarin like?

Everyone who speaks a language has a sense of what is typical and atypical in the pronunciation of their language. This sensitivity reflects their broad experience of the language. For instance, they may notice that their local speech community sounds somewhat different from that in another area, and perhaps none of these local speech varieties sound like the 'standard' TV news anchor. Still, all of these groups are recognized as native speakers and their different pronunciations are within the realm of what is typical. Very loosely then, a foreign-accent is pronunciation that is outside of the typical range, not just of the local speech community, but of the broader community recognized as native speakers of the language.

There are many ways pronunciation can differ across accents. Pronunciation that directly affects words in Mandarin includes segments (vowels and consonants) and suprasegments (tones and perhaps stress). Other aspects of accent create impressions across phrases or longer stretches of speech. These include the rhythm, intonation, speech rate, and pauses that speakers produce. For the moment, we will focus specifically on segmental and tonal speech sounds.

2.1 The Speech Sound Distributions of a Language

When we think about vowels, consonants, and tones, we usually have a specific list—or inventory—of sounds in mind. This inventory includes all the sound categories that make up our words and sentences. Although we can give these categories labels (for example, the /m/ and /a/ sounds in *ma*), the truth is that whenever we produce one of these sounds, it's never exactly the same as the time before. This is true for a single speaker, and is certainly the case across speakers. Our different body shapes and sizes, and our different linguistic experiences all lead to large variability in the sounds we produce. Although we recognize patterns in the pronunciation of our language, there is actually great variability under the surface.

This is illustrated in Figure 1, where we can picture each individual utterance of a sound as a single point in space. The dimensions of that space (*x* and *y*) will be measurable physical

properties of the sound. For example, its duration and fundamental frequency (F0, which we perceive as pitch), or vowel formants (F1, F2, F3—the energy of vibrations in the air within certain frequency ranges). If we measured many instances of the same sound being uttered, we could form a distribution for that sound category (i.e., what listeners perceive as being the same sound). This distribution will look rather circular, with the most typical instances of the sound accumulating at the center of the shape, and less typical instances spreading out towards or beyond the edges. With enough instances and enough different speakers, our circular shape will be a reasonable representation of the typical values of that speech sound.

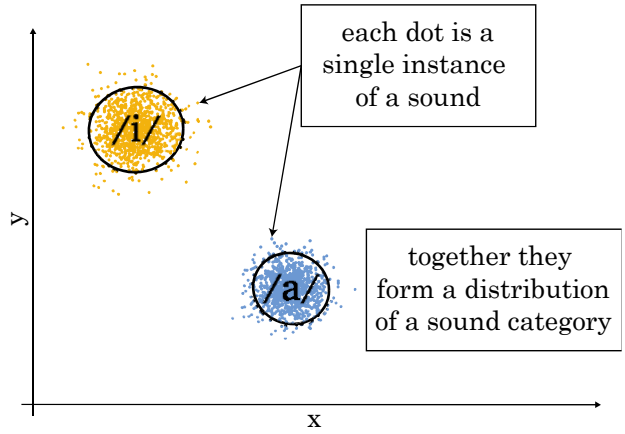

Figure 1. A visualization of speech sound distributions. The *x* and *y* axis represent two separate acoustic measurements such as two vowel formants, or pitch and duration.

This visualization can help us think about accented pronunciation. In simple terms, when sounds fall outside the distribution of typical values, they are accented. If most or all of the sounds a speaker produces fall outside the typical range, that speaker will be perceived to have an accent. (For a much more thorough and technical description of these issues, see Kleinschmidt and Jaeger 2015).

To make this description more concrete, let's consider an example from L2 Mandarin. Figure 2 shows what a hypothetical American English speaker's /a/ sound might look like when they produce Mandarin. By comparing values of the vowel formants (F1 and F2), we can see how similar or different the distribution of the Mandarin /a/ sound (Pinyin *a*) is when produced by our imaginary L1 and an L2 speakers. Notice, the L2 vowel distribution slightly overlaps with the L1 distribution, indicating that sometimes the L2 vowel sounds nativelike.

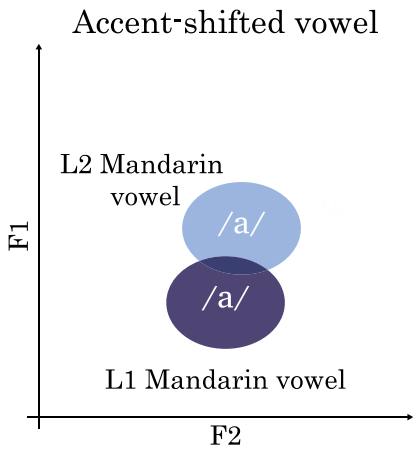

Figure 2. Illustration of accent-shifted L1 Mandarin vowel

2.2 Accent-shifted Pronunciation and Pronunciation Errors

This way of thinking about accented speech gives us the ability to highlight some specific phenomena that often occur in L2 pronunciation. I will describe them as *accent-shifted pronunciation*, and *pronunciation errors*. These are illustrated in Figure 3. (This presentation expands on ideas laid out in Pelzl, Carlson, Guo, Jackson, and van Hell 2020).

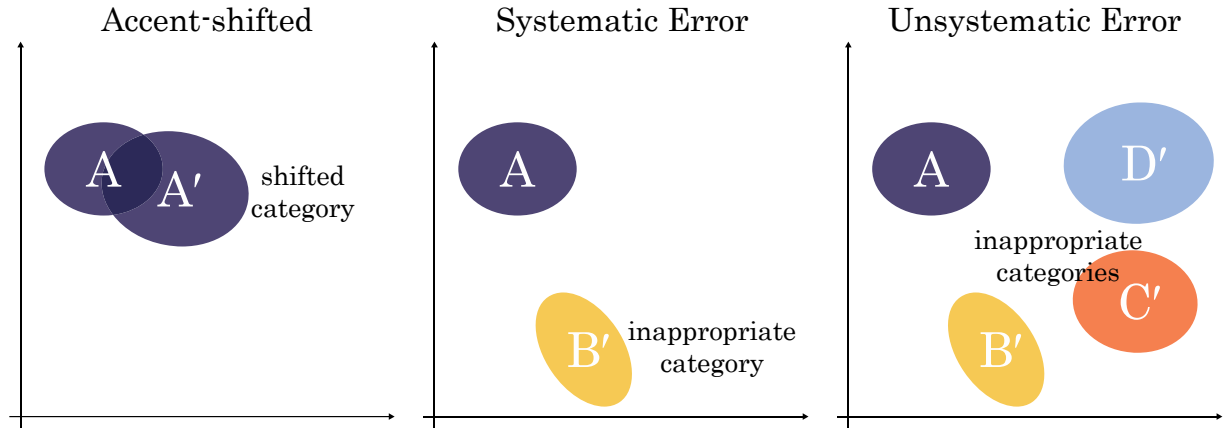

Figure 3. Illustration of types of accent and error. Accent-shifted pronunciation occurs when a speaker produces the intended sound (A) as a shifted (A′) version. Systematic pronunciation error occurs when an inappropriate category (B′) is regularly substituted for the appropriate category (A). Unsystematic pronunciation error occurs when multiple inappropriate categories (B′, C′, D′) are substituted for the appropriate one (A). Figure adapted with permission from Pelzl, Carlson, Guo, Jackson, and van Hell (2020).

2.2.1 Accent-shifted Pronunciation

The left panel in Figure 3 shows the distribution of an accent-shifted pronunciation for speech category A. For now, this could be any sound. The L2 speaker produces their own distribution of the sound (A′), and some instances of it fall within the range of the L1 category

distribution, but most do not. The result is an accent-shifted sound that will often be recognizably different than the typical L1 sound. Importantly, however, the L2 version of the sound is not *randomly* different. There's a pattern that will become clear with enough experience. Presumably, then, listeners will be able to adapt to this type of accented pronunciation. Adaptation in this case means that listeners can learn the new sound pattern, and quickly recognize accented words containing that sound as being those that the L2 speaker intended. Research with foreign-accented English has shown exactly this type of adaptation (e.g., Baese-Berk, Bradlow, and Wright 2013; Bradlow and Bent 2008; Clarke and Garrett 2004; Xie et al. 2018). This does not necessarily mean that listening to foreign-accented speech becomes effortless (McLaughlin and Van Engen 2020), but perhaps with enough experience, it would be.

When it comes to L2 Mandarin, most researchers have not discussed pronunciation in terms of foreign accent, but the type of accent-shifted pronunciation pattern described above is nevertheless documented for a variety of L2 Mandarin consonants and vowels (*consonants:* Hao 2012b; Lai 2009; Liu and Jongman 2013; Shi 2008; X. Wang and Chen 2020; C. Yang and Yu 2019; *vowels:* Hao 2012b; Shi 2009; Wu 2011; Wu and Shih 2012).

Similarly, L2 tones are described in ways that I would call accent-shifted. As this may be a novel way to think about tones, we can consider a few examples. L2 tones have been described as often having an overall F0 range that is constrained compared to L1 tones (G. Chen 1974; Shen 1989; B. Yang 2015: Chapter 7). Beginning learners have been reported to produce the high Tone 1 as either too high or too low (Miracle 1989; Shen 1989; Y. Wang, Jongman, and Sereno 2003), and the pitch onset of Tone 4 has been described as too low relative to the speaker's overall F0 range (Shen 1989; Y. Wang et al. 2003; Zhang 2010). Yang (2015: Chapter 4) also discusses patterns that may be influenced by prosodic factors (intonation, phrasing), observing some L2 speakers to consistently over- or under-shoot tones in certain prosodic locations.

According to the analysis presented here, all of these could be considered accent-shifted versions of tones—rather than tone errors. The accent-shifted L2 tones are different from native patterns, but hypothetically should be recognizable to listeners after they gain some experience with the L2 speaker. However, no research has yet tested this hypothesis.

2.2.2 Pronunciation Errors

In addition to having a foreign accent, another reality for L2 speakers is that they often produce pronunciation errors. In many discussions of foreign accent, errors are simply ignored, or accent-shifted features are described as errors. This is understandable. For listeners, both accents and errors are all wrapped up in the same speech signal and it may not be obvious which sounds are intentional and which accidental. It's also true that a pronunciation error can at the same time be an accent-shifted sound. Still, by drawing sharper distinctions, we can appreciate ways that accent and error from each other, both in terms of why the L2 speaker produces them, and how they might impact listeners.

For L2 speakers, pronunciation errors can be caused by many factors, including inability to hear or form the sounds, insufficient motor muscle control to consistently produce the sounds, or even not knowing what sound is supposed to belong to a given word. Depending on the specific mix of factors, we can outline two broad types of errors: systematic errors that occur with a regular pattern; and unsystematic errors that have no clear pattern.

2.2.3 Systematic Pronunciation Errors

The middle panel in Figure 3 depicts systematic pronunciation errors. In this case the L2 speaker produces a sound (B′) that, for the listener, is categorically different from the typical occurrence of sound A. From the listener's point of view, this is a pronunciation error because it is not the sound they expected to hear. For the speaker, it may well be that they are trying to produce the correct sound, but failing. However, as in the case of accent-shifted sounds, systematic pronunciation errors occur with a pattern. The L2 speaker regularly (if not always) swaps the intended sound with their L2 version of it. In the end then, even though the pronunciation error may be odd, with sufficient experience, a listener could learn the pattern behind it, and adapt so that they more easily and quickly understand the speaker.

As a specific illustration, an L2 speaker of Mandarin may regularly produce the vowel /y/ (as in *l<u>ǜ</u>* 'green') as something closer to what the L1 listener expects to be /u/ (as in *l<u>ù</u>* 'road'). The result would be that words like *lǜ* and *lù* sound the same or much more similar than they should. The pattern does not have to result in another word. For example, an L2 speaker's /p/ (as in *<u>b</u>à* 'dad') could sound like an English speakers /b/ (as in '<u>b</u>ee'). This would not sound quite right, and some listeners might judge it to be an error—but it would also not sound like Mandarin /p^h/ (as in *<u>p</u>à* 'be afraid'). Anecdotally, these examples are actually sounds that English speaking learners of Mandarin struggle to get right. So, though we lack empirical studies about segmental pronunciation errors in L2 Mandarin, we do have reason to suspect that this type of systematic error pattern will often apply to consonants and vowels (for L2 *perception* of /y/, see Hao, 2017).

Systematic errors are also possible, and likely, for tones. It has been suggested that L2 speakers often produce Tone 1 as a falling tone (Miracle 1989; Shen 1989; Y. Wang et al. 2003; C. Yang 2016), and there may also be positional influences that regularly result in tone swaps or distortions of a certain type (*in disyllabic words:* Zhang and Xie 2020; *in phrases:* C. Yang 2016). The pattern behind these positional errors might be more difficult for listeners to learn, but as it is a pattern, there is still a chance they will. For speakers of other tonal languages who learn Mandarin as a second tonal language, there may also be consistent tonal errors that happen due to the influence of the tone categories in their L1. For example, Hao (2012) found Cantonese speakers often swapping the high Tones 1 and falling Tone 4 in their L2 Mandarin productions.

2.2.4 Unsystematic Pronunciation Errors

The final, right-most panel in Figure 3 depicts unsystematic errors. Here the production of the L2 speaker varies so that multiple inappropriate sound categories are produced for what ought to be a single category. There are a few common causes for unsystematic errors. First, the L2 speaker may not be able to perceive the target speech sound, leading to uncertainty about what it ought to sound like. When they need to produce that sound, they simply make a guess or follow some mistaken intuition about what ought to be produced. In this case, the problem is their knowledge of the sound itself. A related problem is that they may forget what sound a word should have, or be mistaken about what they remember. When several words ought to have the same sound, that sound may instead be different for each word. Sometimes the L2 speaker swaps sound A with sound B, sometimes with sound C or D. Finally, the error may be driven by a

physical lack of control. This might occur with sounds in certain positions in a phrase, or due to emotions, or perhaps nervousness. In all cases, the result for the listener is similar—there is an error, but the cause and direction of the error is not clear.

Unlike accent-shifted pronunciation or systematic pronunciation errors, even with extensive experience listeners will not be able to learn the pattern of unsystematic errors, because there is none. If the unsystematic errors happen with enough frequency, listeners may "adapt" in the sense of learning to ignore pronunciation errors. But whereas adaptation to systematic features of L2 speech improves the speed and ease of understanding the speaker, this type of negative adaptation would only serve to remove a source of interference, pushing the listener to rely more heavily on other contextual cues. This might not actually lead to more efficient or easier comprehension of the L2 speaker. Given that a lifetime of experience has taught listeners to automatically use pronunciation for word recognition, it may be the case that they cannot actually learn to ignore unsystematic pronunciation errors.

Unsystematic errors affecting consonants and vowels may not be common. This is partly because these sounds tend to have simple two-way distinctions, so any category swaps would naturally lead to a discernable pattern. A speaker who mispronounces the Mandarin /p/ is likely to waver between /b/ and /p^h/, but not to produce /k/. One instance where it may apply in L2 Mandarin is with the high-rounded front vowel /y/ (Pinyin *ü*), mentioned above. The systematic swapping between /y/ and /u/ could be further complicated if the speaker sometimes also produced the sound as /i/ (as in *lì* 'force'). If this happened with no discernable pattern, it would be an unsystematic error.

What about tones? Here it is not only conceivable, but likely quite common for unsystematic errors to occur. Figure 4 recasts the earlier figure to depict how accent-shifted pronunciation and pronunciation error types might apply to tones. In just a moment, we will consider this in much greater detail.

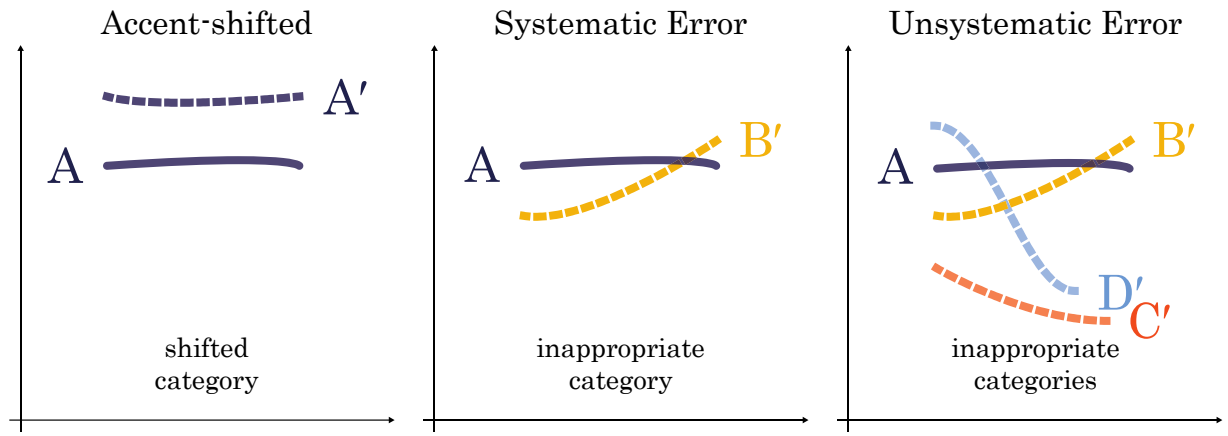

Figure 4. Illustration of instances of tone accent and tone error. Accent-shifted tones occur when a speaker produces the intended tone (A) as a shifted (A′) version. Systematic tone error occurs when an inappropriate tone category (B′) is regularly substituted for the appropriate category (A). Unsystematic tone error occurs when multiple inappropriate categories (B′, C′, D′) are substituted for the appropriate one (A).

2.2.5 Out of Inventory Errors

One additional error pattern that is worth highlighting, especially when thinking about tones, is the *out of inventory* error (Figure 5). That is, an L2 tone category that simply doesn't exist in the language. This type of error could be either systematic or unsystematic, and in some cases may be just an extreme form of an accent-shifted tone. For example, a beginning L2 speaker might sometimes produce a high tone that is shifted so high as to be judged no longer within the conceivable boundaries a well-formed high tone.

Zhang (2010) reports that approximately 14% of tone errors made by her L2 participants were judged as out of inventory by raters, and that these tones were mainly realized as a mid-tone or a low-falling tone. Zhang (2010) is somewhat unique in providing this type of analysis. Most studies have not commented on whether errors are in or out of the Mandarin tone inventory.

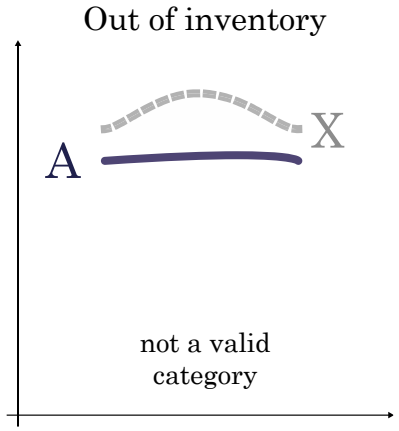

Figure 5. Illustration of a tone error that is not just an inappropriate category, but a non-existent category (X).

2.3 How Frequent are Tone Errors in L2 Speech?

Returning to unsystematic pronunciation errors, the importance of this distinction for tones will depend heavily on whether this type of error is frequent in L2 speech. We do not yet have a clear answer, but there are reasons to suspect they are quite frequent. In research with my colleagues (Pelzl, Lau, Guo, and DeKeyser in press), we found that advanced L2 speakers often have gaps in tone knowledge for about 20% of the words they otherwise know confidently. That is, they know the meanings, but not the tones. As these learners know thousands of words, this suggests they will make tone errors for hundreds or even thousands of specific words. From what we can see so far, there appears to be little pattern to what words L2 speakers do or do not know the tones for. It is not the case, for example, that one specific tone is always the culprit, or that errors are always a switch of the same two tones. This lack of clear patterns would seem to make the occurrence and direction of these lexically-based L2 tone errors largely unpredictable for listeners.

Unfortunately, we can't do much more than speculate at this point. As far as I know, the distinctions drawn here (and in Pelzl et al. 2020) are novel, and so no studies have attempted to characterize the accentedness of L2 tones, or to diagnose whether tone errors are systematic or unsystematic. Still, it may be useful to do a short survey of L2 tone production studies to get a sense for how common tone errors (of any type) are, and why we may or may not have noticed the presence of unsystematic errors in earlier studies.

Among beginning L2 Mandarin speakers, research suggest tone errors may be *very* frequent. Chen, Wee, Tong, Ma, and Li (2016) created a large corpus of beginning L2 Mandarin speech, with speakers from a wide variety of L1s. They report that tone errors occurred on 32% of all syllables produced. It's worth stressing that this was for the reading of isolated syllables with tones explicitly marked in Pinyin. Explicit notation of tones with the Pinyin diacritics (ā á ǎ à) eliminates the memory component from tone production, and provides an iconic cue to the pitch contour. In other words, elicitation using Pinyin is likely to decrease the occurrence of unsystematic errors.

For more experienced speakers, we certainly expect that the frequency of tone errors will be lower. Estimating based on information available in C. Yang (2016: Chapter 3), third and fourth year students seemed—on average—to make errors on about 10% of syllables in a reading passage of about 200 characters. Hao (2012) doesn't provide an overall error rate, but it can be seen that for a reading task, errors of some types (e.g., swapping T3 with T2) occurred nearly 30% of the time. Once again, both of these results are for reading with tones explicitly provided.

For spontaneous L2 Mandarin speech, error frequency may be more difficult to judge. Two studies used relatively unscripted responses to question prompts (Kim et al. 2015; Winke 2007). Both report greater than 90% overall tone accuracy. Considering results from more controlled elicitation methods, this is a rather striking finding. Though both studies report a high degree of consistency between raters, it still may be the case that different approaches to training raters would have increased the detection of errors. The fact that the spontaneously elicited speech in these studies was contextualized may also have reduced raters' sensitivity to pronunciation errors. In any case, even 90% accuracy would still mean a speaker makes an error on one in ten syllables.

2.4 How to Investigate Tone Errors in Future Research

While there will never be a single answer as to the best approach to eliciting L2 speech, the elicitation method is never neutral, and will directly impact what we find (e.g., Hao and de Jong 2016). Reading tasks with tones annotations are often favored because they give us a large degree of control over the specific tonal patterns that speakers (attempt to) produce. They may also be a good method for determining how much control L2 speakers have in ideal circumstances. However, these reading tasks will not tell us about a speaker's knowledge of tones for words. Free or planned responses to question prompts may give a better sense of the frequency and type of tone errors that occur, but it can be very difficult to elicit specific words or tone sequences in such tasks.

To date, most L2 tone production research has been framed around questions of the relative difficulty of the different tone categories. Future work might also attempt to analyze the systematic or unsystematic nature of those errors. This will require the use of elicitation methods other than reading tasks, in order to give unsystematic errors a chance to occur. Some general approaches might include describing pictures or using question-answer pairs that strongly guide the form of the elicited speech. L2 spoken language corpora would be a potentially invaluable resource for understanding broad trends across L2 speakers. In the opposite direction, targeted studies of individual learners who, impressionistically, produce many or few tone errors could also provide insight into the individual differences that lead to L2 tone errors.

What applies to tone errors is also true for segmental speech errors, which have rarely been examined in L2 Mandarin (but see N. F. Chen et al. 2016). Though the unsystematic error type is less likely to occur for segments, it may be that systematic features of segmental L2 speech could influence or be influenced by the frequency and type of errors that occur for L2 tones. For example, perhaps unsystematic tone errors will force listeners to rely more heavily on segmental aspects of L2 speech, thus increasing the importance of clear pronunciation for those segmental features.

In summary, foreign-accented Mandarin includes accent-shifted pronunciation and pronunciation errors. These features are typical of L2 speech, though their frequency will vary from speaker to speaker. A key question for teachers is: how important is it for L2 learners to overcome accent and reduce errors? The next section begins to address this question.

3 How does Foreign-accentedness Affect the Comprehensibility of L2 Mandarin?

People who learn a new language as adults overwhelmingly speak with some degree of foreign accent (e.g., Flege et al. 1995). This does not mean that they cannot improve their pronunciation, but it does raise an important question. *Is a foreign accent a barrier to communication?* Without even conducting any research, we know that extreme answers will not be correct. Widely shared experience tells us that it is not necessary for an L2 speaker to sound exactly like a native speaker in order to communicate effectively. On the other hand, there are certainly cases where a foreign accent can create communication difficulties.

Research in L2 pronunciation has built on these intuitions by trying to measure the relationship between an L2 speaker's accentedness and the comprehensibility of their speech. A highly cited study by Munro and Derwing (1995) suggests the relationship may not be particularly strong. When asked to rate L2 speech samples for accentedness (from weak to strong) and comprehensibility (how easy or difficult a listener finds the speech to understand), they found that even speech rated as strongly accented could still also be rated as highly comprehensible.

These results focused on English. So, as we turn our gaze to Chinese and other tonal languages, a reasonable first question is whether this key finding—that heavily accented speech can also be highly comprehensible—holds for foreign-accented Mandarin speech? Unsurprisingly, from the very start researchers have also wanted to know how tones fit into this relationship.

3.1 The Relationship between Accentedness and Comprehensibility in L2 Mandarin

Lee and Xing (2012) were the first to directly investigate these questions in Mandarin. To explore how tones and segments impacted accentedness and comprehensibility ratings, they made recordings of native Korean L2 speakers of Mandarin reading five simple sentences (e.g., *Jīnwǎn kěnéng huì xiàyǔ.* "It might rain tonight."). Native Mandarin speakers also produced the same five sentences. Lee and Xing then synthesized versions of the sentences with the prosody (i.e., intonation and tones) and segmental features swapped, so that there were sentences with Korean L2 segments and L1 Mandarin prosody, as well as sentences with Korean L2 prosody and L1 Mandarin segments. These manipulated sentences were then rated by a group of native

Chinese listeners. Results showed a clear difference in the perceived accentedness of the manipulated sentences. When L2 segments were present (with L1 prosody), accent was rated more strongly than when L2 prosody was present (with L1 segments). The authors interpret this as evidence that L2 segmental features in Mandarin are more important in conveying accentedness than are the prosodic features—which, of course, includes tones. Like Munro and Derwing (1995), comprehensibility ratings failed to show a strong relationship with accentedness ratings. However, this could be because the sentences were very simple and repeated many times over the course of the study, so that comprehensibility was never a serious issue for listeners after they had heard the sentences a few times. It should also be noted that a single statistical significance test cannot provide support for the absence of an effect.

Lee and Xing's study is the only one to date that has attempted to make a direct comparison between segmental and tonal (prosodic) features of foreign-accented Mandarin. The result is striking, and might suggest tones are not as important as segmental pronunciation in L2 Mandarin. Unfortunately, there are some missing details that make it difficult to fully evaluate the outcomes. Specifically, we do not know what the L2 speakers' tones were like in the recorded stimuli. Were they accent-shifted tones? Did they include tone errors? Accented but otherwise accurate tones might not be expected to have much impact on ratings, whereas outright tone errors would be expected to have much stronger impacts. The small number of very simple stimulus sentences also raises some questions about the generalizability of results to more complicated and varied L2 speech. What happens when vocabulary is not so frequent and predictable? Nevertheless, Lee and Xing applied an interesting approach that might be worth pursuing further in future work.

Working with native English speaking Mandarin learners, Yang also evaluated the role of tones and prosody in foreign-accented speech (C. Yang 2016: Chapter 8). A group of native Chinese raters listened to a small number of short sentences read by either L1 or L2 Mandarin speakers. The raters transcribed the sentences, rated the comprehensibility and accentedness of the speaker, and provided some indication of what they had based their ratings on. Results suggested strong correlations between the accuracy of transcriptions and the ratings of comprehensibility and accentedness—that is, the stronger a speaker's accent, the less comprehensible listeners thought that speaker was.

On its face, this contrasts with the results in English (Derwing and Munro 1995). However, Yang's stimulus sentences were quite different from those used in previous accent studies. Whereas those studies typically had people describe pictures or read narrative passages, Yang's sentences were crafted with much more specific features in mind. Each sentence was exactly six syllables long, had tightly controlled tone patterns, avoided many of the Mandarin consonants, and always contained a rather tricky word-boundary ambiguity (C. Yang 2016: Chapter 4, pp. 60-61). For example, the sentence "*Wū Ānyīng xiū feījī.*" ("Wu Anying repairs planes.") has only the high Tone 1, and requires (like all sentences did) a subtle difference in prosodic phrasing in order to disambiguate whether the proper name is two or three syllables long. With the change of just one written character (and slightly different phrasing), the sentence becomes "*Wū Ān **yīng** xiū feījī.* " ("Wu An should repair planes."). These tricky sentences resulted in a rather large number of transcription errors even when they were produced by native speakers. These challenging stimuli contributed heavily to the outcomes. We can certainly conclude that accent *can* contribute to comprehensibility—and likely will when prosodic or tone ambiguities are present. We cannot conclude that it *usually* does so, because spoken language

typically occurs in context and quite rarely has either the tonal or prosodic features seen in these stimuli.

Freeborn and Rogers (2019) also carried out a rating study with foreign-accented Mandarin, though their aims were a bit different. They wanted to establish whether individual differences among learners would relate to ratings of accentedness. Four L1 and seventy L2 Mandarin speakers—from a wide variety of language backgrounds—read a passage in Chinese with Pinyin annotations. Fifteen L1 Mandarin listeners rated the accentedness of the first two sentences produced by each speaker. Using these ratings, Freeborn and Rogers explored how a large set of seventeen different speaker variables were related to the ratings. Variables included things like current age, age when the speaker began learning, musical training, and so on. The strongest relationship to L1 ratings turned out to be the L2 speaker's own rating of their personal accentedness, with proficiency level (participant's level on the *Hanyu Shuiping Kaoshi*, a standard test of Chinese proficiency used in the PRC), and motivation as the second and third most related factors.

The authors argue that these results show the importance of tones for L2 accentedness. However, this interpretation is not very convincing. Their study had no objective measure of tones at all. Their arguments are based on speculation that L2 learners' ratings of their own accentedness depended on their experience of having conversational breakdowns caused by poor control of tones. This chain of logic might be correct, but they provide little evidence to support it. Additionally, there are reasons to be skeptical of the statistical outcomes in the study given the large number of variables and relatively small number of ratings.

Though not a full-blown rating study, a follow-up question for participants in a study I conducted with my colleagues may also shed some light on the question of tone and accentedness (Pelzl et al. 2020). As shown in Figure 6, we found that L1 Chinese listeners consistently judged an L2 speaker as more accented when that speaker produced tone errors compared to when that same speaker did not produce tone errors. This suggests that tone errors do play some role in producing impressions of a foreign accent. However, this result does not necessarily show special importance for tones over other aspects of L2 pronunciation. Our study had only two L2 speakers, they produced only isolated disyllabic words, and the study design specifically contrasted speakers with respect to their control of tones. Just like in Yang's study (2016), these factors were likely to make tones (and tone errors) highly salient. (See also the chapter by Kaidi Chen and Chunsheng Yang in this volume.)

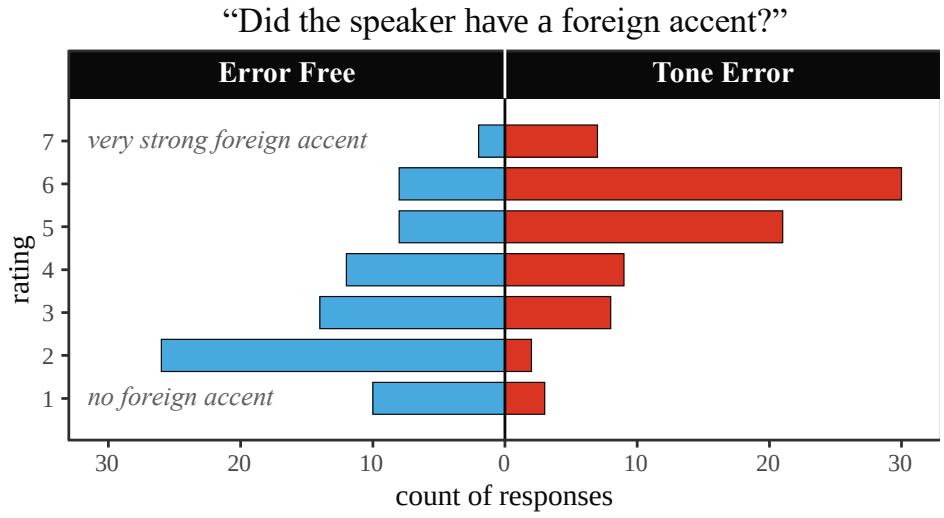

Figure 6. Listener ratings of foreign-accent strength when the same L2 speakers either produced Mandarin without tone errors (Error Free), or with frequent tone errors (Tone Error). Figure adapted with permission from Pelzl, Carlson, Guo, Jackson, and van Hell (2020).

Though not specifically investigating foreign-accented Mandarin, researchers who work with hearing or speech impaired populations also want to understand how tones affect the comprehensibility of Mandarin speech. A number of studies have examined the role of tones by flattening or otherwise manipulating the F0 contours of words and sentences, and then having listeners perform transcription or rating tasks with those sentences. Patel, Xu, and Wang (2010) presented sentences with either their original tones, or a monotone across the whole sentence. In quiet background, the monotone sentences did not cause difficulty for listeners. However, when multi-speaker babble noise was added, listeners were less accurate in transcribing the monotone sentences compared to sentences with tones intact. Further research has shown that flattened tones may have even stronger impacts on elderly or hearing impaired Mandarin listeners (Jiang, Li, Shu, Zhang, and Zhang 2017). These lines of work suggest that similar difficulties would be likely for L2 speech, where tones are not just flattened, but often misleading. Speech-in-noise research could be valuable for understanding how tones in foreign-accented speech might interact with natural (noisy) environments.

Once again, we should exercise some caution when interpreting these studies. On the one hand, simplistic interpretation of results could lead us to underestimate the value of tones. In these studies, no direct contrast was made with segmental features, so results only speak to the impact of tones when segmental pronunciation is accurate. Listeners might rely on tones more heavily when segmental pronunciation is less clear. At the same time, a simplistic reading of results could exaggerate the importance of tones. The stimuli sentences were designed to be challenging and to test comprehension. For example, Patel and colleagues used relatively formal news language, which may present different lexical challenges than most typical L2 speech. Other studies have used word lists, or nonsense sentences, specifically aiming to remove the benefits of meaningful context. Of course, context matters (J. Wang, Shu, Zhang, Liu, and Zhang 2013). When listeners can rely on context, they may be able to easily overcome some of the challenges that misleading tones (or foreign-accented speech) might otherwise present.

In summary, current research clearly shows that tones *can* be a marker of L2 accent, and that in adverse listening conditions or when words are ambiguous because of tones, they can contribute to difficulties in comprehension. However, if pitted against segmental features, it remains unclear whether tones play an equal, greater, or lesser general role in creating the impression of accent or interfering in smooth comprehension.

3.2 Can Listeners Adapt to Foreign-accented Mandarin?

Even when listeners initially find foreign-accented speech difficult to comprehend, we know they can often adapt. People can improve in word recognition for specific accented sounds after hearing just a handful of sentences (Clarke and Garrett 2004; Xie et al. 2018). They get better at transcription of foreign-accented speech over time, regardless of the strength of a speaker's accent (Bradlow and Bent 2008; for a review, see Baese-Berk, McLaughlin, and McGowan 2020).

These positive trends are encouraging. However, this is an instance where the differences between accent-shifted pronunciation and pronunciation errors may become quite important.

When accented speech has a pattern, listeners should be able to adapt. When errors undermine the presence of an obvious pattern, listeners may be unable to adapt, or perhaps will adapt by ignoring the errors and looking elsewhere to guide comprehension. This latter outcome is one possible interpretation for a number of neuro-imaging studies that have found listeners displaying different brain responses to grammatical and lexical errors if those errors are produced by foreign-accented speakers rather than native speakers (e.g., Caffarra and Martin 2018; Grey and van Hell 2017; Hanulíková, van Alphen, van Goch, and Weber 2012; Romero-Rivas, Martin, and Costa 2015).

Taking cues from such studies, my colleagues and I used behavioral and neural measures to test how L1 Mandarin listeners responded to pronunciation errors that occurred in spoken sentences (Pelzl, Lau, Guo, Jackson, and Gor in press). Two speakers read a large number of sentences. One was a native speaker with a typical (Beijing) Mandarin accent, the other was an American L2 speaker of Mandarin. Listeners heard the sentences while their electroencephalogram (EEG, 'brainwaves') was recorded, and for each sentence they judged whether or not they had heard a word or pronunciation error. We wanted to know whether the listeners would respond differently to tonal and segmental pronunciation errors depending on which speaker produced them. The behavioral judgments of listeners made it clear that they responded differently to the foreign-accented speaker—some listeners seemed to find errors even in his 'good' sentences. At the same time, as a group, listeners were more likely to judge sentences with tone errors as acceptable if they were produced by the L2 speaker. This may mean they ignored or did not notice some of the L2 tone errors, but it could also indicate they had more difficulty judging tone errors in foreign-accented speech compared to native speech.

Listeners' neural responses did not show any major differences between the two speakers. There was, however, an overall trend that fits with previous accent studies, indicating that perhaps listeners are less likely to be surprised or even to notice fine-grained pronunciation errors from a foreign-accented speaker. This trend was similar for tonal and segmental pronunciation errors.

Unfortunately, we did not gather more information about *why* listeners made the judgments they did. It could be that some had more or less experience hearing foreign-accented Mandarin (e.g., on TV, among friends), had different levels of strictness in deciding what an error was, or focused on different aims during the task (i.e. comprehending the message vs. judging pronunciation). We also did not get ratings of accentedness or comprehensibility during this study.

Finally, even though one of our goals was to investigate adaptation to foreign-accented Mandarin, we did not find any evidence of changes in listener responses over the course of the study. Failure to find adaptation, however, does not mean adaptation did not occur. Perhaps a different task or response, or simply a larger scale study (more participants) would find evidence of adaptation. It is also possible that the way we manipulated pronunciation (both tonal and segmental) resulted in arbitrary and unsystematic errors. As argued above, it may be impossible for listeners to adapt to this type of error.

In another study (Pelzl et al. 2020), we focused in more narrowly on tones, specifically aiming to examine the effects of unsystematic tone errors. Two L2 speakers—we'll call them speaker A and speaker B—produced isolated two-syllable words. On each trial in the study, native Mandarin listeners heard an L2 speaker produce a word and then saw a written Chinese word. In some cases, the written words matched what was spoken; in others, the written word was different. This was meant to create a *priming* effect so that responses would be faster when

words were the same in both spoken and written form. Native Chinese participants all heard both L2 speakers, but for half of the participants speaker A made tone errors on 50% of filler words, while speaker B made no tone errors. For the other half of participants, this was reversed: speaker B made 50% tone errors, speaker A made none. Our question was whether the frequency of tone errors would slow down listeners' recognition of words, even when the words were produced correctly. If so, this would be a strong argument for the negative effects of unsystematic tone errors on L2 comprehension.

The answer from this single study was negative. It didn't matter whether or not the L2 speaker made tone errors, listeners always responded equally fast when words were spoken correctly. At the same time, when tone errors did occur, listeners were a bit slower to recognize the words. For example, if they heard the incorrectly produced ***nènglì*** and then saw the 'matching' written word 能力, they were slightly slower to recognize it as a real word. In short, we found that tone errors have a direct impact on the speed with which listeners recognize words, but we did not find any evidence that listeners adapted to a speaker who made frequent tone errors.

Practically speaking, the results of our two studies show that, for two-syllable words, tone errors do impact the speed and efficiency with which listeners recognize words both in isolation and in context. At a minimum then, tone errors seem likely to increase the effort needed to understand foreign-accented Mandarin. We cannot be sure the same patterns would apply for single syllable words, where tone errors are much more likely to result in a completely different words, rather than merely mispronounced ones. This would suggest single syllable words will lead to more confusion—but it has to be balanced against the fact that many single syllable words are extremely frequent in conversation (Tao 2015), and likely to be easily inferred in context.

4. Future Directions for Foreign-accented Mandarin Research

As research on foreign-accented Mandarin is just beginning, there are many basic questions that can be asked. For those interested in research with practical applications to classrooms, I will take a moment to consider three of the main questions whose answers might provide significant guidance for teaching practices.

4.1 What Specific Sounds may be Most Important to Target in Pronunciation Teaching?

Given the major role tones play in L2 pedagogy and the challenge they present to many learners, the focus on tones in current research is understandable. Another reason that tones may be a popular topic of study is that, with only a handful of them, it is much more tractable to target them all at once, compared to consonants and vowels. Still, whatever the ultimate findings are for the importance of tones, it will not mean that consonants and vowels don't matter.

Future research might try to find a route into segments by evaluating whether some consonants or vowels are more important than others. In research on English, one interesting approach to this question has been through the lens of *functional load* (Kang and Moran 2014; Munro and Derwing 2006; Suzukida and Saito 2019). Essentially, the idea is that some sets of contrastive sounds may be more important than others, because—across the spoken

vocabulary—they serve to distinguish more words. For example, /b/ and /p/ in English differentiate many words (bit/pit, back/pack, bat/pat, etc.) and thus have a high functional load. In contrast, the sounds /θ/ and /ð/ (as in '**th**igh' and '**th**y') differentiate very few words and so have a low functional load. Though so far somewhat exploratory, the studies that have investigated these issues in English seem promising. For work along these lines in Mandarin, guidance can be sought from a very active line of research addressing the informational and statistical properties of consonants, vowel, and tones (Tong, Francis, and Gandour 2008; Wiener 2020; Wiener and Ito 2015, 2016; Wiener, Lee, and Tao 2019; Wiener and Turnbull 2015; Yao and Sharma 2017).

Additionally, existing studies on L2 tone production can guide explorations about how specific tonal features impact listeners' perceptions of accentedness or the actual comprehension of L2 speech. For example, recent discussions about the best approach to teaching Tone 3 might gain further clarity by gathering listener responses to L2 speech (e.g., He, Wang, and Wayland 2016; J. Shi 2007; Sparvoli 2017; Wen and Yan 2015; Zhang 2014).

4.2 How do Prosodic Features of Foreign-accented Speech Impact Comprehensibility?

Tones and segments are not the only important aspects of pronunciation. In English language research, suprasegmental aspects of foreign-accented speech—intonation, stress, speech rate—have received quite a bit of attention (Kang 2010; Munro 1995). Some studies have suggested training on those features does more to increase L2 comprehensibility than training only on segmental features (Derwing, Munro, and Wiebe 1998; Derwing and Rossiter 2003). Future work in Mandarin would do well to also consider these prosodic features of foreign-accented speech. As mentioned above, this was one part of Yang's (2016) study, and Lee and Xing (2012) also describe their study in terms of prosody, rather than just tones. By expanding from this work, and also incorporating insights from other L2 research, we can begin to test whether broader prosodic trends might deserve more attention in Chinese classroom teaching.

4.3 What are the Social Implications of Foreign-Accented Mandarin?

Even when foreign-accented pronunciation does not impede comprehensibility, it often comes with social costs. I began this chapter by referencing the phrase *yáng qiāng yáng diào*, which is used to refer to the speech of foreign-accented Mandarin speakers. Though the specific implications of the phrase can be shaped by many contextual factors, it often bears a negative connotation (DeFrancis 2003). So then, whether we like it or not, it is worth understanding the social costs associated with foreign-accented speech, as well as what L2 speakers can or cannot do to mitigate those costs.

As mentioned briefly above, one type of social cost comes from the increased effort foreign-accented speech sometimes requires of listeners (McLaughlin and Van Engen 2020). Not every person will have the same amount patience and determination when communicating with an L2 speaker. Insofar as L2 speakers can improve their pronunciation, they may be able to lessen the burden on their listeners.

Unfortunately, not every social cost can be mitigated by improved L2 pronunciation. Social psychologists have found bias towards or against individuals based on their appearance,

such that the same vocal recordings presented with different faces resulted in different judgments of accentedness—a phenomenon that has come to be called "reverse linguistic stereotyping" (Kang and Rubin 2009, 2014; Rubin 1992). Undoubtedly, similar things occur among Chinese listeners who may be biased to expect foreign-accented Mandarin from those who fit their expectations of what L2 speakers look like (i.e., non-Chinese), or alternatively, biased to expect nativelikeness from those who look like L1 speakers (i.e., appear Chinese). Research in these areas should be conducted with due sensitivity, but could be very useful for understanding what is and isn't in the control of the L2 speaker.

Relatedly, additional work could be conducted looking at the role of non-standard (regional) Chinese accents when produced by L2 speakers. Diao (2017) has conducted one interesting study along these lines, considering L2 speakers who chose to retain regional features in their Mandarin speech.

For all research into foreign-accent, it will of course be important to determine what results are broadly generalizable across different native language groups, and what results are more dependent on the L2 speaker's specific linguistic experience.

5. Conclusion

Every speaker has an accent. L2 speakers of Mandarin are no different. By studying the ways that foreign-accented speech affects listeners, we can slowly build towards a more empirically driven understanding of what needs to happen for learners to communicate effectively in Mandarin. This work is just beginning, I hope that in another ten years, a review like this will have more numerous and more concrete results to share.

Acknowledgements: This research was supported in part by the National Science Foundation (SBE 2004279).

References

Baese-Berk, M. M., Bradlow, A. R., & Wright, B. A. (2013). Accent-independent adaptation to foreign accented speech. *The Journal of the Acoustical Society of America*, *133*(3), EL174–EL180. https://doi.org/10.1121/1.4789864

Baese-Berk, M. M., McLaughlin, D. J., & McGowan, K. B. (2020). Perception of non-native speech. *Language and Linguistics Compass*, *14*(7). https://doi.org/10.1111/lnc3.12375

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*(2), 707–729. https://doi.org/10.1016/j.cognition.2007.04.005

Caffarra, S., & Martin, C. D. (2019). Not all errors are the same: ERP sensitivity to error typicality in foreign accented speech perception. *Cortex*, *116*, 308–320. https://doi.org/10.1016/j.cortex.2018.03.007

Chen, G. (1974). The pitch range of English and Chinese speakers. *Journal of Chinese Linguistics*, *2*, 159–171.

Chen, N. F., Wee, D., Tong, R., Ma, B., & Li, H. (2016). Large-scale characterization of non-native Mandarin Chinese spoken by speakers of European origin: Analysis on iCALL. *Speech Communication*, *84*, 46–56. https://doi.org/10.1016/j.specom.2016.07.005

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *The Journal of the Acoustical Society of America*, *116*(6), 3647. https://doi.org/10.1121/1.1815131

DeFrancis, J. (2003). *ABC Chinese-English comprehensive dictionary: Alphabetically based computerized*. Honolulu: University of Hawai'i Press.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in Favor of a Broad Framework for Pronunciation Instruction. *Language Learning*, *48*(3), 393–410. https://doi.org/10.1111/0023-8333.00047

Derwing, T. M., & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning*, *13*(1), 1–17.

Diao, W. (2017). Between the standard and non-standard: Accent and identity among transnational Mandarin speakers studying abroad in China. *System*, *71*, 87–101. https://doi.org/10.1016/j.system.2017.09.013

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*(04). https://doi.org/10.1017/S0272263104040021

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 233–277). Timonium, MD: York.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, *97*(5), 3125–3134.

Freeborn, L., & Rogers, J. (2019). Nonlinguistic Factors that Affect the Degree of Foreign Accent in Second Language Mandarin. *Studies in Chinese Linguistics*, *40*(1), 75–99. https://doi.org/10.2478/scl-2019-0003

Grey, S., & van Hell, J. G. (2017). Foreign-accented speaker identity affects neural correlates of language comprehension. *Journal of Neurolinguistics*, *42*, 93–108. https://doi.org/10.1016/j.jneuroling.2016.12.001

Hanulíková, A., van Alphen, P. M., van Goch, M. M., & Weber, A. (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, *24*(4), 878–887.

Hao, Y.-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. Journal of Phonetics, *40*(2), 269-279. https://doi.org/10.1016/j.wocn.2011.11.001

Hao, Y.-C. (2018). Second Language Perception of Mandarin Vowels and Tones. *Language and Speech*, *61*(1):135-152. https://doi.org/10.1177/0023830917717759

Hao, Y.-C., & de Jong, K. (2016). Imitation of second language sounds in relation to L2 perception and production. *Journal of Phonetics*, *54*, 151–168. https://doi.org/10.1016/j.wocn.2015.10.003

He, Y., Wang, Q., & Wayland, R. (2016). Effects of different teaching methods on the production of Mandarin tone 3 by English speaking learners. *Chinese as a Second Language*, *51*(3), 252–265.

Jiang, W., Li, Y., Shu, H., Zhang, L., & Zhang, Y. (2017). Use of semantic context and F0 contours by older listeners during Mandarin speech recognition in quiet and single-talker interference conditions. *The Journal of the Acoustical Society of America*, *141*(4), EL338–EL344. https://doi.org/10.1121/1.4979565

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, *38*(2), 301–315. https://doi.org/10.1016/j.system.2010.01.005

Kang, O., & Moran, M. (2014). Functional loads of pronunciation features in nonnative speakers' oral assessment. *TESOL Quarterly*, *48*(1), 176–187. https://doi.org/10.1002/tesq.152

Kang, O., & Rubin, D. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology*, *28*(4), 441–456.

Kang, O., & Rubin, D. (2014). Reverse linguistics stereotyping. In J. Levis & A. Moyer (Eds.), *Social dynamics in second language acquisition* (pp. 239–253). Berlin: DeGruyter Mouton.

Kim, J., Dewey, D. P., Baker-Smemoe, W., Ring, S., Westover, A., & Eggett, D. L. (2015). L2 development during study abroad in China. *System*, *55*, 123–133. https://doi.org/10.1016/j.system.2015.10.005

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148–203. https://doi.org/10.1037/a0038695

Lai, Y. (2009). Asymmetry in Mandarin affricate perception by learners of Mandarin Chinese. *Language and Cognitive Processes*, *24*(7–8), 1265–1285. https://doi.org/10.1080/01690960802113850

Lee, J.-K., & Xing, L. (2012). The relative weight of prosody and segment in the perception of Korean accented Chinese speech. *Language and Linguistics (Korean Journal)*, *54*, 263–293.

Liu, J., & Jongman, A. (2013). *American Chinese learners' acquisition of L2 Chinese affricates /ts/ and /tsh/*. 060005–060005. Kansas City, Missouri. https://doi.org/10.1121/1.4798223

Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. New York, NY: Routledge.

McLaughlin, D. J., & Van Engen, K. J. (2020). Task-evoked pupil response for accurately recognized accented speech. *The Journal of the Acoustical Society of America*, *147*(2), EL151–EL156. https://doi.org/10.1121/10.0000718

Miracle, W. C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of the Chinese Language Teachers Association*, *24*(3), 49–65.

Munro, M. J. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition*, *17*, 17–34.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech and second language learners. *Language Learning*, *45*(1), 73–97.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, *34*(4), 520–531. https://doi.org/10.1016/j.system.2006.09.004

Norman, J. (1988). *Chinese*. New York, New York: Cambridge University Press.

Patel, A. D., Xu, Y., & Wang, B. (2010). The role of F0 variation in the intelligibility of Mandarin sentences. *Speech Prosody*, 4.

Pelzl, E., Carlson, M. T., Guo, T., Jackson, C. N., & van Hell, J. G. (2020). Tuning out tone errors? Native listeners do not down-weight tones when hearing unsystematic tone errors in foreign-accented Mandarin. *Bilingualism: Language and Cognition*, 1–8. https://doi.org/10.1017/S1366728920000280

Pelzl, E., Lau, E. F., Guo, T., & DeKeyser, R. (in press). Even in the best-case scenario L2 learners have pesistent difficulty perceiving and utilizing tones in Mandarin: Findings from behavioral and event-related potentials experiments. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S027226312000039X

Pelzl, E., Lau, E. F., Guo, T., Jackson, S. R., & Gor, K. (in press). Behavioral and neural responses to tone errors in foreign-accented Mandarin. *Language Learning*. https://doi.org/10.1111/lang.12438

Romero-Rivas, C., Martin, C. D., & Costa, A. (2015). Processing changes when listening to foreign-accented speech. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00167

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*(4), 511–531. https://doi.org/10.1007/BF00973770

Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Teachers Association*, *24*(3), 27–47.

Shi, F. (2008). *The structure of speech sounds: The intersection of phonetics and phonology. 语音格局—语音学与音系学的交汇点*. Beijing, China: Shangwu Yin Shu Guan.

Shi, F. (2009). *Shiyan Yinxixue Tansuo (Exploration of Experimental Phonology)*. Beijing, China: Peking University Press.

Shi, J. (2007). On teaching tone three in Mandarin. *Journal of the Chinese Language Teachers Association*, *42*(2), 1–10.

Sparvoli, C. (2017). From phonological studies to teaching Mandarin tone: Some perspectives on the revision of the tonal inventory. In I. Kecskes & Sun, Chaofen (Eds.), *Key Issues in Chinese as a Second Language Research* (pp. 81–100). New York, NY: Routledge.

Suzukida, Y., & Saito, K. (2019). Which segmental features matter for successful L2 comprehensibility? Revisiting and generalizing the pedagogical value of the functional load principle. *Language Teaching Research*, 136216881985824. https://doi.org/10.1177/1362168819858246

Tao, H. (2015). Profiling the Mandarin Spoken vocabulary based on corpora. In W. S.-Y. Wang & C. Sun (Eds.), *The Oxford Handbook of Chinese Linguistics*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199856336.013.0031

Tong, Y., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes*, *23*(5), 689–708. https://doi.org/10.1080/01690960701728261

Wang, J., Shu, H., Zhang, L., Liu, Z., & Zhang, Y. (2013). The roles of fundamental frequency contours and sentence context in Mandarin Chinese speech intelligibility. *The Journal of*

the Acoustical Society of America, *134*(1), EL91–EL97. https://doi.org/10.1121/1.4811159
Wang, X., & Chen, J. (2020). The acquisition of Mandarin consonants by English learners: The relationship between perception and production. *Languages*, *5*(2), 20. https://doi.org/10.3390/languages5020020
Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America*, *113*(2), 1033. https://doi.org/10.1121/1.1531176
Wen, B., & Yan, F. (2015). Liuxuesheng Hanyu shengdiao xide zhong yang-shang heliude shiyan yanjiu. *Journal of the Chinese Language Teachers Association*, *50*(1), 19–41.
Wiener, S. (2020). Second language learners develop non-native lexical processing biases. *Bilingualism: Language and Cognition*, *23*(1), 119–130. https://doi.org/10.1017/S1366728918001165
Wiener, S., & Ito, K. (2015). Do syllable-specific tonal probabilities guide lexical access? Evidence from Mandarin, Shanghai and Cantonese speakers. *Language, Cognition and Neuroscience*, *30*(9), 1048–1060. https://doi.org/10.1080/23273798.2014.946934
Wiener, S., & Ito, K. (2016). Impoverished acoustic input triggers probability-based tone processing in mono-dialectal Mandarin listeners. *Journal of Phonetics*, *56*, 38–51. https://doi.org/10.1016/j.wocn.2016.02.001
Wiener, S., Lee, C., & Tao, L. (2019). Statistical regularities affect the perception of second language speech: Evidence from adult classroom learners of Mandarin Chinese. *Language Learning*, *69*(3), 527–558. https://doi.org/10.1111/lang.12342
Wiener, S., & Turnbull, R. (2015). Constraints of tones, vowels and consonants on lexical selection in Mandarin Chinese. *Language and Speech*, 0023830915578000.
Winke, P. M. (2007). Tuning into tones: The effect of L1 background on L2 Chinese learners' tonal production. *Journal of the Chinese Language Teachers Association*, *42*(3), 21–55.
Wu, C. (2011). *The evaluation of second language fluency and foreign accent*. University of Illinois at Urbana-Champaign.
Wu, C., & Shih, C. (2012). *A corpus study of native and non-native vowel quality*. Presented at the Speech Prosody, Shanghai, China.
Xie, X., Weatherholtz, K., Bainton, L., Rowe, E., Burchill, Z., Liu, L., & Jaeger, T. F. (2018). Rapid adaptation to foreign-accented speech and its transfer to an unfamiliar talker. *The Journal of the Acoustical Society of America*, *143*(4), 2013–2031. https://doi.org/10.1121/1.5027410
Yang, B. (2015). *Perception and Production of Mandarin Tones by Native Speakers and L2 Learners*. New York: Springer.
Yang, C. (2016). *The Acquisition of L2 Mandarin Prosody: From experimental studies to pedagogical practice*. Philadelphia, PA: John Benjamins Publishing Co.
Yang, C., & Yu, A. C. L. (2019). The acquisition of Mandarin affricates by American second language learners. *Taiwan Journal of Linguistics*, *17*(2), 91–122. https://doi.org/10.6519/TJL.201907_17(2).0004
Yao, Y., & Sharma, B. (2017). What is in the neighborhood of a tonal syllable? Evidence from auditory lexical decision in Mandarin Chinese. *Proceedings of the Linguistic Society of America*, *2*, 45. https://doi.org/10.3765/plsa.v2i0.4090
Zhang, H. (2010). Phonological universals and tonal acquisition. *Journal of the Chinese Language Teachers Association*, *45*(1), 39–65.

Zhang, H. (2014). The third tone: Allophones, sandhi rules and pedagogy. *Journal of the Chinese Language Teachers Association*, *49*(1), 117–145.

Zhang, H., & Xie, Y. (2020). Coarticulation effects of contour tones in second language Chinese. *Chinese as a Second Language Research*, *9*(1), 1–30. https://doi.org/10.1515/caslar-2020-0001