*The phonological determinants of tone in English loanwords in Mandarin**

Eleanor Glewwe iD
Grinnell College

This paper presents the results of a corpus study and an online loanword adaptation experiment examining the tonal adaptation of English loanwords in Mandarin. Using maximum entropy models, I control for the substantial influences of lexical tone distributions and standardisation, and uncover phonological determinants of tone beyond these lexical and conventional factors. The most important phonological determinant of tone in the corpus was English voicing, while in the experiment it was English stress-aligned pitch contours. I argue that these distinct tonal adaptation patterns constitute two different perceptual mappings, one from F0 perturbations to tone and the other from English intonation to tone, both arising due to particular borrowing contexts. I suggest that increasingly close contact between English and Mandarin may lead to more intonation-driven tonal adaptation in the latest wave of borrowing. The maximum entropy approach holds promise for the analysis of complex cases of tonal adaptation in other languages.

1 Introduction

Loanword phonology is a robust field of study that has yielded new insights into native phonological grammars and has shed light on cross-linguistic perception. Within loanword phonology, the study of suprasegmental adaptation considers how stress, tone and pitch accent are treated in borrowing. As the literature on suprasegmental adaptation

has grown (e.g. Kang 2010, Davis *et al.* 2012), the case of Mandarin tonal adaptation has stood out as especially problematic. Numerous studies have sought to account for the tones of English loanwords in Mandarin, but nearly all fail to fully explain the considerable variation in tonal assignment. The complexity of Mandarin tonal adaptation is typically attributed to the roles played by non-phonological, even non-linguistic, factors, including the morphosyllabic writing system with which loanwords must be written (Chang 2020).

In the present study, I investigate Mandarin tonal assignment in both established English loanwords and online adaptations. These turn out to behave differently. My study addresses three main questions: (i) what are the determinants of tonal assignment in English loanwords?; (ii) what are the phonological determinants of tone, if any, and how can they be accounted for within models of loanword adaptation?; (iii) on what basis do integrated loanwords and online adaptations differ with respect to tonal adaptation? (i) and (ii) are closely tied, because isolating the phonological determinants of tone requires understanding other sources of loanword tones and because the literature is divided as to how phonological tonal adaptation is. (iii) might be expected to be related to (ii), since past work suggests that online adaptations may involve more direct phonological or perceptual mappings than integrated loanwords (Shinohara 2000, Kubozono 2006). In addressing these three questions, I aim to uncover a more comprehensive picture of Mandarin tonal adaptation than has yet been established.

In the remainder of this section, I review major themes in the tonal adaptation literature, in order to provide context for the English-to-Mandarin case. First, tonal adaptation, like segmental adaptation, is restricted by the phonology of the borrowing language. Native constraints on which tones can occur in which positions or with which syllable shapes may limit tonal assignment in a loanword, whatever the properties of the source form (Osatananda 1996, Kenstowicz & Suchato 2006, Hsieh & Kenstowicz 2008). Even the lexicon can constrain tonal adaptation: in Mandarin, accidental tonotactic gaps may restrict which tone can be assigned to an adapted syllable, since each syllable must be written with an existing Chinese character (Wu 2006).

One recurrent motif in the tonal adaptation literature is the concept of 'stress-to-tone' principles, exemplified by English loanwords in Cantonese (Silverman 1992, Hao 2009). Stress-to-tone occurs when the acoustic correlates of stress in the source language are perceptually similar enough to the lexical tones of the borrowing language for stress to be adapted as tone. In Cantonese, English stress-aligned pitch accents and intonational tones are interpreted as lexical tone sequences: stressed syllables are assigned high tone, pretonic unstressed syllables mid tone and posttonic unstressed syllables low tone (Chen 2000, Hao 2009).[1]

---

[1] Cantonese tonal adaptation exhibits additional patterns beyond the basic stress-to-tone mapping (Silverman 1992, Hao 2009).

Stress-to-tone also arises in the adaptation of English loanwords in Hausa (Leben 1996) and Yoruba (Kenstowicz 2006).

In other cases, tonal adaptation depends not on source language tones or intonation patterns, but on segmental properties. In several tone languages, including Thai (Kenstowicz & Suchato 2006), White Hmong (Golston & Yang 2001) and Taiwanese (Hsieh 2006), tones are assigned to loanwords based at least in part on syllable structure. In assigning tones to Mandarin loanwords, Lhasa Tibetan ignores the original Mandarin tones, despite having tones that could adapt them. Instead, syllables beginning with obstruents, all of which are voiceless in Mandarin, receive high tone, while syllables beginning with sonorants receive low tone (Hsieh & Kenstowicz 2008). Such cases show that some tonal adaptation cannot be explained by mappings from pitch properties in the source language to tones in the borrowing language.

In the present study, a final consideration is the potential for adaptation patterns to differ between established loanwords and online adaptations. An example comes from pitch accent assignment in English loanwords in Japanese. Kubozono (2006) proposes that integrated English loanwords are typically accented, rather than unaccented, because the pitch fall heard in English citation forms (whose intonation pattern is presumably H* L-L%) is perceived as the presence of a pitch accent. The exact location of the pitch drop in English is not preserved in Japanese, though; instead, a native phonological rule assigns the default accentuation pattern. In contrast, Shinohara (2000) found that pitch accent assignment preserved the location of stress in Japanese speakers' online adaptations of English words. What underlies a difference in behaviour between established loanwords and elicited adaptations may vary from case to case. De Jong & Cho (2012) note that integrated loanwords involve a degree of time depth that online adaptations do not, and may therefore exhibit patterns of abstraction and regularisation. Online adaptations, on the other hand, may be influenced by the immediacy of contact and fixed form of the source language stimulus that are characteristic of the laboratory setting.

For the present study, I conducted a corpus study of established English loanwords in Mandarin, and an English-to-Mandarin loanword adaptation experiment.[2] In the corpus study, tonal adaptation was mostly non-phonological, but the strongest phonological determinant of tone was voicing in English. Phonological effects played a greater role in the experiment, where the principal phonological determinant of tone was English intonation, driven by stress. I argue that these results reflect two different perceptual mappings, each arising from a specific borrowing context. I then discuss the implications of this study for tonal adaptation in Mandarin and beyond.

---

[2] An earlier analysis of an earlier version of the corpus is given in Glewwe (2016).

2 Tonal adaptation in Mandarin

Before reviewing past work on tonal assignment in English loanwords in Mandarin, I provide some necessary background on both languages. The consonant inventory of Mandarin is given in (1a) (adapted from Lee & Zee 2003, Miao 2005 and Dong 2012), with Pinyin symbols following the IPA representations. (Hereafter, I use Pinyin to transcribe Mandarin.) Aspiration is contrastive in Mandarin stops and affricates (e.g. /p/ *b vs.* /pʰ/ *p*). Voicing, however, is not; all Mandarin obstruents are voiceless.

(1) a. *Mandarin consonant inventory*

	bi-labial	labio-dental	dental	post-alveolar	alveolo-palatal	palatal	velar
stop	p *b*		t *d*				k *g*
	pʰ *p*		tʰ *t*				kʰ *k*
fricative		f *f*	s *s*	ʂ *sh*	ɕ *x*		x *h*
affricate			ts *z*	tʂ *zh*	tɕ *j*		
			tsʰ *c*	tʂʰ *ch*	tɕʰ *q*		
nasal	m *m*		n *n*				ŋ *ng*
approxi-mant	w *w*		l *l*	ɻ̣ *r*		j *y*	

b. *English consonant inventory*

	bi-labial	labio-dental	inter-dental	alve-olar	post-alveolar	pal-atal	velar	glot-tal
stop	p b			t d			k g	
fricative		f v	θ ð	s z	ʃ ʒ			h
affricate					tʃ dʒ			
nasal	m			n			ŋ	
approxi-mant	w			l	ɹ	j	(w)	

The consonant inventory of RP English, which is used in Dong's (2012) corpus, is given in (1b) (based on Collins & Mees 1996). Unlike in Mandarin, voicing is contrastive in English stops, fricatives and affricates. While English voiceless stops and affricates have aspirated and unaspirated allophones, aspiration is not contrastive. However, since voiced stops and affricates are often devoiced in positions where voiceless stops and affricates are aspirated (i.e. word-initially and initially in a stressed syllable), aspiration is arguably contrastive in these positions.

Miao (2005) identifies the general correspondence between English and Mandarin consonants in segmental adaptation. Her findings largely hold for my corpus. Most relevantly, for stops and affricates, the English voicing contrast is mapped onto the Mandarin aspiration contrast. No

such contrast-preserving correspondence is possible for English voiced and voiceless fricatives, however, so both are mapped to Mandarin voiceless fricatives.

Mandarin has four lexical tones, exemplified in (2). The second column shows how the tones are represented in Pinyin (note that Pinyin tone diacritics differ from IPA notation). Additionally, Mandarin has a fifth 'neutral' tone (also analysed as the absence or loss of tone), which appears on some grammatical morphemes and the final syllables of some compounds, and whose realisation is context-dependent (Duanmu 2007).[3]

(2) *Tone*   *Pinyin*
    1 (high)   mā   'mother'
    2 (rising)   má   'hemp'
    3 (low)   mǎ   'horse'
    4 (falling)   mà   'scold'
    5 (neutral)   ma   (question particle)

Not all Mandarin syllables are found with every tone. While some of these tonotactic gaps are accidental, Wu (2006) points out that others are systematic. In particular, syllables with sonorants onsets often lack high tone pronunciations (e.g. *lān*), and syllables with unaspirated onsets tend to lack rising tone pronunciations (e.g. *dán*).[4] With a handful of historical exceptions, Mandarin does not fill lexical gaps in loanword adaptation, so the only syllables that are available for rendering borrowings are those already attested in the lexicon (Wu 2006).[5]

Previous literature on Mandarin tonal adaptation has presented a number of sometimes conflicting claims about what drives tonal assignment in English borrowings, whether established or online. A difficulty common to most studies is the sheer variability in the tones assigned within a single corpus: while adaptation patterns can be uncovered, much variability often remains unexplained. Here, I extract the main threads and findings from past work.

Wu's (2006) seminal study identifies a number of determinants of tone, including stress, onset consonant and the Mandarin lexicon, that have come up repeatedly in subsequent work. In a corpus of around 100 English loanwords (excluding proper names) in Taiwanese Mandarin, Wu found that English (stressed) monosyllables were adapted with falling tone (e.g. *show → xiù*). The initial stressed syllables of English trochees tended to receive high or rising tone (e.g. *curry → kālǐ*; *modern →*

---

[3] The third tone is realised as a falling–rising tone in pre-pause position (Duanmu 2007).

[4] Syllables that do not occur with rising tone underlyingly can be realised with rising tone as a result of third tone sandhi (i.e. *dǎn* can surface as *dán* before another syllable bearing third tone).

[5] Wang (2020) contends that this lexical conservatism follows from the need to write loanwords in Chinese characters, and provides examples of modern borrowings that are not written in characters and that fill tonotactic and even phonotactic gaps.

módēng) (a preference confirmed experimentally by Mar & Park 2012), but there were many exceptions. These patterns, though not absolute, could instantiate stress-to-tone insofar as falling tone best matches the citation H* L-L% intonation pattern realised on monosyllables, and high and rising tones at least partially match the H* aligned with initial stressed syllables. Other studies have also reported stress-to-tone effects. Chang & Bradley's (2012) adaptation experiment found a falling tone preference for all stressed syllables, while Zheng & Durvasula's (2016) experiment found that English–Mandarin bilinguals most often assigned high tone to initial stressed syllables. In Li's (2017) corpus study, stressed syllables rarely received low tone.

Wu (2006) examined the tonal adaptation of initial stressed syllables in disyllabic loanwords more closely, concluding that the choice between high and rising tone depended on properties of the onset of these syllables in Mandarin. Syllables beginning with a sonorant in the adapted form tended to bear rising tone, while syllables beginning with obstruents tended to bear high tone. Within obstruents, aspiration also played a role: adapted syllables with aspirated onsets were more likely to have rising tone, while those with unaspirated onsets were more likely to have high tone. Wu attributed these segmental patterns to depressor effects of sonorant and aspirated onsets on F0 and to the systematic tonotactic gaps in the lexicon discussed above, both of which give rise to associations between certain types of onsets and particular tones that then guide the behaviour of adapters. Wu also conducted an experiment in which Taiwanese Mandarin speakers had to choose between high and rising tone adaptations of the initial syllables of English nonce trochees. The results corroborated the corpus findings, though the aspirated-rising tone pattern was weaker in the experiment. Later studies finding Mandarin onset effects consistent with Wu's include Chang & Bradley's (2012) and Mar & Park's (2012) experiments and Zheng & Durvasula's (2016) corpus study of personal names. In a previous corpus study, Glewwe (2016) uncovered effects of *English* consonants: loanword syllables whose onsets corresponded to English sonorants and voiced obstruents preferred rising tone, while syllables whose onsets corresponded to English voiceless obstruents preferred high tone.

Since the literature has paid substantial attention to sonorants and aspirated stops as depressor consonants, it is worth considering the evidence for their depressor effects. Such evidence is needed to argue for a perceptual mapping from F0 perturbations to tone (though this is not necessarily the claim in previous work). While the lowering of F0 after voiced stops (*vs.* voiceless stops) is robustly attested cross-linguistically (Kingston & Diehl 1994), sonorant onsets do not have a consistent effect on F0, and there is debate as to whether they can be depressor consonants (Maddieson 1997, Kingston 2011). Hombert *et al.* (1979) reported lowering of F0 after sonorant onsets in English, but they believed this was an experiment-specific intonational effect, not a true lowering effect. Li (2005) found that initial sonorants had no perturbation effect on F0 in

Mandarin. The effect of aspiration on F0 is also unclear. Cross-linguistically, the reported effects are inconsistent (Hombert 1976, Xu & Xu 2003). In English, aspirated and unaspirated voiceless stops pattern together, with higher F0 after both types of voiceless stop than after voiced stops (Ohde 1984), while in Mandarin, aspiration lowers F0 at the onset of voicing (Xu & Xu 2003). F0 perturbations induced by aspirated Mandarin onsets thus might underlie the segmental effect identified by Wu, but the picture is less certain for sonorant onsets.

Wu notes one further pattern in initial stressed syllables: four out of eight English syllables with voiced stop onsets were adapted with aspirated stop onsets (*vs.* the expected unaspirated stop onsets) *and* rising tone. She proposes that the depressor effect of English voiced stops on F0 sometimes leads to an adaptation with rising tone, which in turn leads to an adaptation with an aspirated stop, due to the Mandarin-internal aspiration–rising tone association. However, her claim is based on very few syllables.

Before moving on from onset effects, it is also worth briefly considering the matter of source language *vs.* borrowing language segments. Previous literature has tended to focus on the segmental properties of Mandarin, particularly aspiration. The reason for this is not always clear, though it may have to do with the assumption that tonal adaptation is subordinate to segmental adaptation (Wu 2006, Chang 2020). In any case, this approach supposes that onset effects in tonal adaptation are driven not by properties of English consonants (at least, not directly), but rather by Mandarin-internal tone–segment interactions.[6] The choice between English and Mandarin onsets is less important in the case of sonorants, for instance, since English sonorants are virtually always adapted as Mandarin sonorants. If English aspiration was significant, however, this would be hard to discover from an examination of Mandarin onsets alone, because segmental adaptation does not distinguish English aspirated and unaspirated voiceless stops. The potential effects of English segmental properties on Mandarin tonal assignment have thus been somewhat neglected in past research.

While Wu touches on the role of the Mandarin lexicon in tonal adaptation, later work has pushed this idea much further. Zheng & Durvasula (2016) hypothesise that the tone choices of Mandarin monolinguals in their experiment reflects the distribution of lexical tones in Mandarin. This factor is further investigated in Chang (2020). Chang's major claims are that tonal adaptation takes place after segmental adaptation, and that it is primarily non-phonological. She reanalysed Chang & Bradley's experiment, concluding that participants assigned tones to adapted syllables by following native tone probabilities. For example, the probabilities of a given token of *pa* bearing high, rising, low and falling tones in running Mandarin text are 0.05, 0.17, 0.00 and 0.78

---

[6] Alternatively, English consonants may influence tonal adaptation, but segmental and tonal adaptations interact with each other according to Mandarin-internal tone–segment patterns (see Wu's 2006 hypothesis for why some English syllables with voiced stop onsets are adapted with aspirated onsets and rising tone).

respectively. If participants base their tone choices on lexical tone frequencies, their tonal adaptations of any loanword syllable *pa* should reflect this distribution. Chang argues that these tone probabilities accounted for most tonal assignments in the experiment, with stress on English syllables perhaps boosting high tone assignment and reducing low tone assignment slightly (Chang analyses only the adaptations of stressed syllables). She also constructed two corpora of place names: a formal dictionary corpus and an informal blog corpus. Again, native tone distribution predicted most tonal assignment. Chang attributes most deviations from lexical tone probabilities to character avoidance, i.e. deliberate attempts to avoid certain high-frequency characters. She ascribes other deviations to institutionally imposed conventions applied in the dictionary corpus. Zheng & Durvasula (2016) similarly acknowledge that adaptation in their corpus of names may have been partially standardised, and Mar & Park (2012) recognise that there exists a set of characters commonly used in loanwords.

Summing up, the previous literature presents a complex picture of Mandarin tonal adaptation. Proposed determinants of tonal assignment include stress-to-tone, onset consonants, Mandarin lexical tone frequencies, character avoidance, and convention; of these, stress-to-tone and onset effects are phonological in nature. Studies disagree as to the relative weighting of these factors, which usually cannot even explain all the observed variation in a given set of loanwords. The picture is also mixed when it comes to differences between established loanwords and online adaptations. For instance, while Wu (2006) found similar onset effects in her corpus and her experiment, Zheng & Durvasula (2016) found that tonal assignment patterns differed between their corpus and the online adaptations they elicited, with the latter apparently reflecting a more surface-level perceptual mapping.

Past corpus studies are small (Wu 2006) or disregard stress (Zheng & Durvasula 2016, Chang 2020), while experiments vary widely in methodology and exclusively consider stressed syllables. In this paper, I analyse a larger corpus and a new experiment, examining both stressed and unstressed syllables. I also control for lexical tone distributions and the influence of standardisation in both established loanwords and online adaptations, in order to isolate and compare phonological determinants of tone in each type of borrowing.

3 Corpus study

3.1 The corpus

To investigate tonal adaptation in established English loanwords in Mandarin, I carried out a new corpus study. The loanwords were drawn mostly from Dong's (2012) dissertation on segmental adaptation. Assembled from four dictionaries, Dong's corpus consisted of about

1200 borrowings, including place names and first names. I supplemented this corpus with words from Wu (2006) and with entries labelled 'loanword' in the online Chinese dictionary *MDBG* (2014).

I excluded about 65 loanwords whose adaptations I presumed were influenced by semantics (e.g. 蹦极 *bèngjí* 'bungee (jumping)', from 蹦 *bèng* 'jump' + 极 *jí* 'extreme'), since the choice of a particular character with an appropriate meaning entails the assignment of a particular tone (see Miao 2005 for further discussion). I also excluded 60 syllables subject to third tone sandhi (Chen 2000), because it is impossible to tell whether it is their surface rising tone or their underlying low tone that reflects how tones are assigned to loanwords (e.g. the first syllable in *fǎlǎo* 'pharaoh'). Additionally, I excluded about 200 tokens of the syllables *a*, *fo*, *hei*, *le*, *miu*, *ri*, *se*, *sen*, *te* and *teng*, because, due to tonotactic gaps, they can only bear one tone in Mandarin. Since the tones on these syllables are fixed, it makes little sense to include them in discussions of tonal assignment.

Another issue complicating the study of Mandarin tonal adaptation is indirect borrowing. Li (2017) and Wang (2020) point out that many English loanwords, especially those borrowed before the 1950s, entered Mandarin via Cantonese or Shanghainese, which have different tonal systems (as well as phonemic inventories and phonotactics). Such loanwords do not bear on Mandarin adaptation patterns, because their pronunciations are simply the Mandarin readings of the written forms of the original Cantonese or Shanghainese adaptations, which would exhibit the segmental and tonal adaptation patterns specific to those languages. To reduce this source of noise as much as possible, I excluded about 25 loanwords that were likely borrowed into Mandarin through Cantonese (cf. Dong 2012). Some appeared in the Cantonese loanword adaptation literature, while I suspected that others had arrived through Cantonese because of the Cantonese form's greater segmental resemblance to the English form (e.g. 威化 *wēihuà* 'wafer'; Cantonese: [wɐi HH faː MM]). I further excluded 啤(酒) *pí(jiǔ)* 'beer', which Wang (2020: 22) identifies as a 'notable Shanghainese loanword'.

Finally, I excluded about 750 epenthetic syllables. Epenthetic syllables (e.g. the second syllable in *āsīpǐlín* 'aspirin') are inserted in adapted forms to preserve consonants from English clusters and codas without violating Mandarin phonotactics. Since epenthetic syllables do not correspond to English syllable nuclei, I do not consider them in the analysis.

After all these exclusions, the corpus consisted of 2644 syllables, from 1287 loanwords.

3.2 Standard syllables in the corpus

Past studies have observed that loanwords in Mandarin show the influence of convention and standardisation (Mar & Park 2012, Zheng & Durvasula 2016, Chang 2020, Wang 2020). Chinese institutions like the Xinhua News Agency issue official adaptations or dictate standards for adapting

foreign words, including proper names, which contributed 72% of the syllables in my corpus (Dong 2012). Since the use of a standard character entails the assignment of a particular tone, and since the considerations that went into choosing these characters are obscure, it is necessary to control for standard characters in studies of tonal adaptation. In this section, I examine the extent to which standard characters and syllables are present in my corpus.

To do this, I consulted the English-to-Mandarin transcription chart in *Names of the world's peoples* (Xinhua News Agency 1993). The chart's columns correspond to onset consonants and its rows to vowel nuclei or syllable rhymes. (The chart does not take stress into account, except insofar as English vowel quality is confounded with stress.) Each cell gives the character prescribed for the adaptation of one or more English syllables (e.g. 巴 *bā* for /bɑː/, /bæ/ or /bʌ/).

I first established what English string each syllable corresponded to (e.g. *pǐ* in *āsīpǐlín* adapts the string /pə/ from *aspirin* /ˈæ.spə.ɹɪn/). I then coded each syllable for whether it used the standard character for the English string it adapted. In this first pass, I erred on the side of allowing standard character use to account for as much of the corpus as possible. While the chart maps English pronunciations to Mandarin syllables, English orthography often appeared to influence which cell a standard character was chosen from. For example, while the standard character for /bə/ is 伯 *bó*, the string /bə/ in *baroque* /bəˈɹɒk/ is adapted with 巴 *bā*, the standard character for /bɑː/. It seems likely that this is because the syllable /bə/ is spelled <ba>, so I coded this syllable, and others like it, as using the standard character. Additionally, if an English string was adapted with the standard character for a different but phonetically very similar string (e.g. /ən/ for /ɪn/), I also coded the loanword syllable as using the standard character.

Applying this coding scheme, I found that 1852, or about 70%, of the syllables in the corpus used standard characters. An additional 99 syllables used a character homophonous with the standard character for the English string they adapted (e.g. the second syllable in *toffee* /ˈtɒ.fi/ is adapted with 妃 *fēi*, homophonous with the standard character 菲 *fēi*). Another 362 syllables were adapted with some standard syllable from the chart, though not the one prescribed for the given English string. For instance, the second syllable in the adaptation of *Martina* /ˌmɑːˈtiː.nə/ is 丁 *dīng*, rather than the 廷 *tíng* prescribed for English /tin/. However, *dīng is* a standard syllable; it is used to adapt English /din/. In the end, only 331 syllables (13%) did not match some standard syllable, showing how pervasive the influence of standardisation was in the corpus.

The use of standard *syllables*, regardless of Chinese characters or English strings, presents a problem for uncovering patterns of tonal adaptation in the corpus. In loanwords like *toffee*, with, as noted above, 妃 *fēi* instead of the standard 菲 *fēi*, adapters may have been aware of the standard loanword syllable *fēi*, and used a different character with the same pronunciation, either due to ignorance of which character was officially prescribed or

for other reasons.[7] The adaptation has nevertheless been shaped by the standard syllable *fēi*. In loanwords like *Martina*, even though adapters used a different segmental syllable than is standard for the given English string, they may still have been guided by their knowledge of the repertoire of standard syllables for loanwords. That is, even if they chose *dīng* instead of the expected *tíng* for *Martina*, this does not mean the adaptation is free of the influence of standardisation: they may have opted for *dīng* because it is also a standard syllable. Since the focus of this study is tonal adaptation, I do not seek to explain the segmental adaptation of the loanwords. If segmental adaptation happens first (Chang 2020), then any time the subsequent step of tonal assignment produces a standard syllable, there is no reason to conclude that the choice of tone indicates anything beyond adapters' awareness of these standard syllables. All standard syllables in the corpus therefore potentially have tones that principally reflect official standards, and reveal little about possible phonological determinants of tone.

The pervasiveness of standard syllables in the loanword corpus constitutes a major confound that threatens to distort any analysis of tonal adaptation patterns. It has also likely contributed to the difficulty that past studies have had in identifying anything more than general trends in the data and in accounting for all the variation (cf. Wang 2020). One way to avoid this confound would be to analyse only the 331 non-standard syllables. Excluding this much data drastically reduces the size of the corpus, though, and having fewer data points would make it harder to test for a wide range of potential phonological influences on tonal assignment.

Another possible pitfall of excluding all the standard syllables is that, despite undoubtedly reflecting official standards, they may still, in the aggregate, reflect some phonological patterning as well. For instance, suppose there were a tendency for adaptations of English stressed syllables to receive high tone. One standard syllable with high tone is *dīng*. The segmental syllable *ding* appears in the corpus as the adaptation of a number of English strings. It could be the case that, beyond *ding*'s 'baseline' tendency to have high tone because the standard syllable is *dīng*, it is even *more* likely to have high tone when the English syllable it corresponds to is stressed. If standard syllables are excluded entirely, patterning like this would be inaccessible to the analysis. Phonological determinants of tone might go undetected, or their strength could be over- or underestimated. Ultimately, I kept the standard syllables in the corpus and carried out an analysis using maximum entropy Harmonic Grammar that controlled for lexical effects and the use of standard syllables.

---

[7] Conceivably, these may include semantic reasons. While I tried to exclude adaptations where the character choices seemed clearly guided by semantics, it is possible that subtler semantic influences remain. This may be a source of persistent noise in corpora of English loanwords in Mandarin.

3.3 MaxEnt analysis of the corpus

Maximum entropy Harmonic Grammar (MaxEnt; Goldwater & Johnson 2003, Hayes & Wilson 2008) is a probabilistic constraint-based framework of phonology that can handle variation in the output for a given input. A MaxEnt grammar has weighted constraints, and candidates for each input receive harmony scores, which are the sums of their weighted constraint violations. The harmony scores are used to generate probabilities of occurrence for each candidate.[8] The most harmonic candidate (i.e. the one with the lowest harmony score) will be the most frequent output, the next most harmonic candidate will be the next most frequent output, and so on. A successful MaxEnt model generates candidate probabilities that closely resemble candidates' observed frequencies. Given a set of inputs, candidates, candidate frequencies, constraints and constraint violations, a learning algorithm can determine the constraint weights that maximise the probability of the observed outputs (i.e. that yield the best match between the candidate frequencies predicted by the MaxEnt grammar and the actual candidate frequencies). Importantly, the MaxEnt models I present are intended not as grammatical models of Mandarin tonal adaptation, but as statistical analyses that allow us to control for lexical tone frequencies and the use of standard syllables, and to test for phonological effects in tonal assignment. Potential phonological determinants of tone are implemented as constraints in a MaxEnt grammar, and the importance accorded to these constraints by the learning algorithm reveals how much influence various phonological factors exert.

I carried out the MaxEnt analysis in three steps. The first step established a model that predicted the tones of loanword syllables based solely on tone distributions in the lexicon. The second step added constraints that captured the additional effects of standard syllables on tonal assignment. In the third step, I examined the role of phonological factors involved in tonal adaptation in the corpus. This three-step analysis represents the most conservative approach to identifying phonological effects in loanword tones. By first modelling the roles of the lexicon and of standard syllables, I attribute as much of the variation in tonal assignment as possible to these factors, so that the third step only uncovers phonological effects that could not be otherwise explained by lexical statistics or reliance on standard syllables with potentially incidental tones.

I conducted the MaxEnt analysis on the full corpus of 2644 syllables. The input to the learning algorithm consisted of a large tableau composed of sub-tableaux, each of which corresponded to one Mandarin segmental syllable associated with various properties of the English input. The English inputs were not fully specified English syllables, but rather abstract units that preserved information about English stress, position and onset properties. I distinguished stressed and unstressed syllables,

---

[8] For details of the MaxEnt formula, see Hayes & Wilson (2008).

and final and non-final syllables.[9] There were seven possible onset types: voiceless aspirated stop/affricate, voiceless unaspirated stop/affricate, voiced stop/affricate, voiceless fricative, voiced fricative, sonorant and zero. Sub-tableaux were organised by Mandarin segmental syllable, however, because I assume, following Chang (2020), that tonal adaptation follows segmental adaptation.[10] The same English syllable can be adapted with different Mandarin segmental syllables: for example, final unstressed /kəʊ/ is rendered as *kě* in *rococo*, but as *kòu* in one version of *cocoa*. I am interested not in how /kəʊ/ came to be adapted as *ke* or *kou*, but in how *kě* came to bear low tone and *kòu* falling tone. Thus, despite corresponding to the same English syllable, *kě* and *kòu* are data points in two different sub-tableaux, one for *ke* and one for *kou*.

Additionally, adaptations of different English syllables may be data points in the same sub-tableau if the English syllables are identical after abstraction and adapted with the same Mandarin segmental syllable. For example, non-final stressed syllables (or strings) such as /ˈdæ/ in *Daphne* /ˈdæf.ni/, /ˈdɑː/ in *Darwin* /ˈdɑː.wɪn/ and /ˈdʌ/ in *dozen* /ˈdʌ.zn̩/, all of which are adapted with Mandarin *da*, do not differ on any phonological factor that I consider. They are therefore all included in a single *da* sub-tableau. This sample sub-tableau, representing 13 tokens from the corpus, is given in (3). The input is the Mandarin segmental syllable *da*, with associated information about the English form (i.e. /ˈdæ/, /ˈdɑː/, /ˈdʌ/, etc.), while the candidates are *da*, with each of the four possible lexical tones. The observed counts are taken from the corpus. Other elements of the sub-tableau are explained below.

(3)

da	ob-served counts	lexical tone proba-bilities	*DA1 2.815	*DA2 1.263	*DA3 1.669	*DA4 0	predicted proba-bilities	pre-dicted counts
a. dā	0	0.039	1				0.039	1
b. dá	12	0.185		1			0.185	2
c. dǎ	1	0.123			1		0.123	2
d. dà	0	0.653				1	0.653	8

Finally, there can be multiple sub-tableaux for the same Mandarin segmental syllable. For instance, while the tokens of *da* in (3) adapt non-final stressed syllables with voiced onsets, other tokens of *da* adapt non-final stressed syllables with voiceless onsets (e.g. the second syllable in *Montana* /ˌmɒn.ˈtæ.nə/). These tokens were entered into a separate *da* sub-tableau, to allow testing of the effects of distinct English onset properties on tonal assignment.

[9] Based on evidence that Mandarin speakers treat English unstressed syllables and syllables with any degree of stress differently (Cheng 1968, Li 2017), I collapsed primary stressed and secondary stressed syllables into a single category.

[10] Alternatively, segmental adaptation might simply constrain tonal adaptation (cf. Wu 2006) without preceding it, a view which would also allow for the present approach.

In the first step of the MaxEnt analysis, I established a model that predicted what tones the syllables in the corpus would have if they were projected directly from the lexical tone frequencies for each segmental syllable. I used Chang's (2020) tone probabilities, which were calculated from Da's (2004) character frequencies. The lexical tone probabilities for *da* are shown in (3). The most probable tone for *da* is falling tone: based on all the characters with the pronunciation *da* and their frequencies in Da's corpus of Modern Chinese texts, the probability that a given token of *da* will have falling tone is 0.653.

To establish the baseline lexical model, I created four constraints for each segmental syllable, each of which penalised the occurrence of that syllable with one of the four tones. Each candidate for a given syllable violated the constraint that penalised its tone. For instance, in (3), the candidate *dā* incurs one violation of *DA1. I used the MaxEnt Grammar Tool (Hayes *et al.* 2009) to fit the constraint weights, not to the observed tone counts in the corpus but to the lexical tone probabilities. In the case of *da*, the constraint *DA1 should have the highest weight, because *dā* has the lowest lexical probability, and therefore needs to be suppressed the most by the constraint that penalises it. I used positive weights for the lexical constraints, so as the weight of *DA1 increases, the harmony score (i.e. badness) of *dā* increases, thereby reducing its predicted probability. The fitted weights of the lexical constraints for *da* are given in (3), where *DA1 has the highest weight.

Since there was a constraint for every possible syllable–tone combination, the model was able to achieve essentially a perfect fit to the native tone frequencies: in (3), the predicted probabilities of the candidates exactly match *da*'s lexical tone probabilities. I converted the model's predicted probabilities to predicted counts for the tones in the corpus. For example, the lexical model predicts that eight of the 13 *da* tokens in this sub-tableau should bear falling tone. In fact, none of them did.

The square of the correlation (r^2) between the observed corpus counts and the predicted counts of the baseline lexical model was 0.585. This goodness-of-fit measure is a way of gauging how much of the variation in tonal assignment the model can explain. Lexical tone probabilities evidently influence what tones loanword syllables receive, but they cannot fully account for tonal assignment in the corpus.

In the second step of the MaxEnt analysis, I added standard syllable constraints to the model. For instance, for the segmental syllable *da* there exists the standard syllable *dá*, so I added a constraint DA2, which, as I will explain, rewards the candidate *dá* for matching a standard syllable. In some cases, there was no standard syllable for a segmental syllable in the corpus (e.g. *zhi*), and in other cases there were multiple standard syllables (e.g. *di* has both *dí*, for /diː/ and /dɪ/, and *dì*, for /tiː/ and /tɪ/), giving rise to multiple constraints.

The standard syllable constraints were added to the baseline lexical model, whose already fitted lexical constraint weights were held constant. Thus the learning algorithm could assign weight to the standard syllable

constraints in order to achieve a better fit to the observed tones in the corpus. In other words, adding the standard syllable constraints and determining their weights allowed me to quantify the role of standard syllables in causing assigned tones to deviate from the lexical tone distributions. (4) shows how the sub-tableau in (3) is updated in the standard syllables model with the addition of the standard syllable constraint DA2. The candidate that 'violates' DA2 is *dá*; however, I fitted the standard syllable constraints with *negative* weights, so that a 'violation' of a standard syllable constraint reduces (i.e. improves) the violating candidate's harmony score, thereby increasing its predicted probability. I used negative weights so that a constraint like DA2, which specifies that a *da* token should have rising tone, could simply reward *dá* by reducing its harmony score, rather than having to penalise *dā*, *dǎ* and *dà* by increasing their harmony scores by the same amount. Negative weights allow for a very straightforward interpretation.

(4)	*da*	observed counts	lexical tone probabilities	*DA1 2.815	*DA2 1.263	*DA3 1.669	*DA4 0	DA2 −3.276	predicted probabilities	predicted counts
	a. dā	0	0.039	1					0.007	0
	b. dá	12	0.185		1			1	0.857	11
	c. dǎ	1	0.123			1			0.022	0
	d. dà	0	0.653				1		0.114	2

In the standard syllables model, the lexical constraints retained their previously fitted weights. They were active in the model insofar as their weights contributed to the candidates' harmony scores, but their weights were fixed, modelling the influence of the lexicon. I used Excel Solver to fit negative weights just to the standard syllable constraints, in such a way that the probability of the observed tones from the corpus was maximised. The predicted probabilities of the standard syllables model were generated from the weights of both the lexical constraints (the role of the lexicon) and the standard syllable constraints (the role of standardisation). (4) shows the new predicted probabilities and counts for *da*. Based on all the *da* sub-tableaux in the model, DA2 received a relatively large negative weight, increasing the predicted incidence of *dá* in this sub-tableau from two tokens in the lexical model to eleven in the standard syllables model. This is closer to the actual number of observed *dá* tokens, so the influence of standardisation on tonal assignment appears to be strong for *da*.

The fit between the observed and predicted counts increased greatly in the standard syllables model ($r^2 = 0.952$, up from 0.585 in the lexical model). Factoring in the effect of standard syllables on top of the tone distributions projected from the lexicon accounts for many more of the assigned tones in the corpus. In fact, it appears that lexical tone probabilities and the use of standard syllables explain nearly all tonal adaptation in established loanwords.

In the third step of the MaxEnt analysis, I investigated what further role, if any, phonological factors played in tonal assignment in the corpus. I retained the lexical and standard syllable constraints, with their fixed weights, from the previous model, and added a host of phonological constraints. By fitting weights to these constraints, I could gauge the presence and strength of phonological effects in the corpus.

I devised a fairly exhaustive set of potential phonological effects, based on findings in the literature. For each phonological property, there were four constraints, one for the association of that property with each tone. Just as a standard syllable constraint like DA2 specifies that tokens of *da* should bear rising tone, a phonological constraint like STRESSED1 specifies that loanword syllables that adapt English stressed syllables should bear high tone. Since past studies reported stress-to-tone effects, I included the constraints STRESSED1–4 and UNTRESSED1–4. I also included FINAL1–4 and NON-FINAL1–4, which applied to loanword syllables adapting word-final and non-word-final English syllables, because of Wu's (2006) finding that monosyllables (and thus perhaps all stressed final syllables) were adapted with falling tone. Moreover, I included constraints that could capture interactions of stress and position: STRESSEDNON-FINAL1–4, STRESSEDFINAL1–4, UNTRESSEDNON-FINAL1–4 and UNSTRESSEDFINAL1–4.

I included constraints for onset properties covering English sonority, English voicing and English and Mandarin aspiration. OBS1–4 and SON1–4 applied to loanword syllables whose onsets corresponded to English obstruents and sonorants respectively. VOICED1–4 and VOICELESS1–4 applied analogously, with respect to the phonemic voicing of the English consonant. ASP1–4 and UNASP1–4 applied to loanword syllables whose onsets correspond to English aspirated and unaspirated stops and affricates respectively, while MANDARINASP1–4 and MANDARINUNASP1–4 applied to loanword syllables with aspirated and unaspirated stop and affricate onsets in Mandarin. To account for the possibility that English voiced obstruents might have a phonological effect distinct from that of all obstruents or of all voiced consonants, including sonorants, I also included the constraints VOICEDOBS1–4. Wu (2006) observes that loanword syllables with aspirated onsets adapting English voiced stops have rising tone. I therefore included the constraints VOICED & MANDARINASP1–4 to check for this interaction of English voicing and Mandarin aspiration. Finally, to test for an across-the-board preference for a particular tone in loanword syllables independent of lexical and standard syllable influences (for example, one could imagine high tone serving as a default tone in loanwords), I added the constraints TONE1–4, each of which simply specifies that loanword syllables should have that tone.

Like the weights fitted to the standard syllable constraints, the weights fitted to the phonological constraints were negative. Thus, for example, a constraint like STRESSED1 rewarded a candidate like *dā* in (4) for being the adaptation of an English stressed syllable and having high tone. The larger

the negative weight of STRESSED1, the more the harmony score of *dā* would decrease, resulting in a higher predicted probability of *dā*. If the predicted probability of *dā* did not need to increase to better match the tones observed in the corpus, then STRESSED1 would not accrue a large negative weight. In this way, the fitted weights of the phonological constraints, as well as the degree to which they contribute significantly to the MaxEnt model, reveal which phonological properties are most influential in tonal assignment.

I again used Excel Solver to determine the weights of the 76 phonological constraints that maximised the probability of the actual tones in the corpus. In this phonological model, the predicted tone probabilities are generated from the lexical, standard syllable and phonological constraints. After finding the weights of the phonological constraints, I removed 15 constraints with weights of zero. Next, I pruned the phonological constraint set by successively removing the constraint that contributed least to the model, as determined by a likelihood ratio test, until removing a single constraint made the model's fit significantly worse at the 0.05 level. The final pruned model contained the 23 phonological constraints in Table I.

In the pruned phonological model, the fit between the observed and predicted tone counts improved slightly ($r^2 = 0.967$ *vs.* 0.952 in the standard syllables model). A likelihood ratio test showed that the addition of the 23 phonological constraints significantly improved the model's performance in predicting the observed tones, justifying their inclusion ($\chi^2(23) = 199.517$, $p < 0.001$). Thus, although the fit between observed and predicted tones increased only modestly from the standard syllables model, there were real phonological effects in the corpus.

Table I lists the 23 phonological constraints in the final MaxEnt model and their weights in descending order of Δ log likelihood (i.e. the decrease in log likelihood if the given constraint is removed from the model; the larger its Δ log likelihood, the more a constraint is contributing to the model). Note that constraint weights and Δ log likelihoods do not go hand in hand: a constraint may have a relatively large weight but a relatively small Δ log likelihood, meaning its contribution to the model is relatively small. This can happen if a constraint's effect is limited to only a few sub-tableaux in the model. For each constraint, Table I also gives the rate at which the relevant tone was assigned to loanword syllables with the relevant phonological property by the standard syllables model (before the addition of phonological constraints), by the present phonological model and in the corpus. For instance, in the case of VOICEDOBS2, the standard syllables model predicts that 35.7% of loanword syllables whose onsets correspond to English voiced obstruents will have rising tone; the addition of the phonological constraints increases that rate to 44.4%. The observed rate is 44.9%.

The top three constraints by Δ log likelihood are VOICEDOBS2, VOICED2 and VOICELESS1, all of which involve the voicing and/or sonority of the English consonant to which a loanword syllable's onset corresponds.

| constraint | weight | Δ log likelihood | predicted rate | | observed rate |
			standard syllables model	phonological model	
VOICEDOBS2	−2.425	134.496	35.7%	44.4%	44.9%
VOICED2	−0.799	40.931	38.3%	40.8%	41.1%
VOICELESS1	−0.857	19.627	39.0%	43.1%	42.4%
MANDARINUNASP1	−0.973	18.972	40.2%	39.4%	38.5%
UNASP4	−1.085	17.376	18.1%	17.3%	16.3%
OBS3	−0.806	16.959	16.0%	15.5%	15.3%
TONE4	−0.354	16.682	27.0%	27.1%	27.7%
VOICED1	−0.752	14.272	19.3%	18.5%	18.6%
OBS1	−0.577	13.331	38.6%	38.9%	38.3%
VOICELESS4	−0.582	11.964	23.9%	26.4%	27.5%
SON1	−1.078	10.562	9.3%	10.9%	11.3%
MANDARINUNASP3	−0.975	9.739	11.1%	10.5%	11.0%
VOICELESS3	−0.749	9.647	20.0%	20.2%	19.6%
SON4	−0.305	8.734	31.0%	29.3%	30.5%
MANDARINUNASP4	−0.914	6.960	19.3%	20.8%	20.8%
MANDARINASP2	−0.957	5.781	22.9%	22.0%	20.8%
VOICED3	−0.561	5.771	16.5%	16.7%	15.6%
STRESSED4	−0.242	5.059	26.0%	25.9%	26.7%
TONE2	−0.184	4.680	28.7%	28.0%	28.4%
STRESSED1	−0.252	3.761	28.8%	30.6%	30.6%
STRESSEDNON-FINAL1	−0.211	2.915	29.9%	31.7%	32.4%
FINAL4	−0.260	2.749	27.1%	28.6%	28.9%
SON3	−0.530	1.940	20.0%	21.0%	19.2%

Table I
Phonological constraints in the corpus MaxEnt analysis.

VOICEDOBS2 has a Δ log likelihood that is nearly an order of magnitude larger than any other phonological constraint, as well as by far the largest negative weight. This means that the strongest phonological effect in the corpus was an increased likelihood of rising tone on loanword syllables whose onsets corresponded to English voiced obstruents, even when lexical tone frequencies and the use of standard syllables were controlled for. The Δ log likelihood of the second most important constraint, VOICED2, also sets it apart from the other 21 phonological constraints, since it is twice that of the next constraint. The strength of VOICED2 reveals a more general association between English voicedness, across obstruents and sonorants, and rising tone assignment. The third most important constraint, VOICELESS1, reflects the flip side of the voicing effect: loanword

syllables whose onsets correspond to English voiceless obstruents are more likely to receive high tone.

What about the phonological constraints with progressively smaller Δ log likelihoods? The next two constraints, MANDARINUNASP1 and UNASP4, are not far behind VOICELESS1 in contributing to the model. The MaxEnt model is a very complex system, in which many intersecting constraints interact to adjust the predicted probabilities of the candidates in each sub-tableau. Consequently, a relatively important constraint does not necessarily reflect a true phonological effect. A heuristic means of checking whether a constraint is likely to represent a real phonological effect is to inspect the improvement in predicted tonal assignment from the standard syllables model to the phonological model. VOICEDOBS2, VOICED2 and VOICELESS1 all push up the assignment rate of the given tone for the relevant syllables, and achieve a closer fit to the observed data, suggesting they represent genuine phonological effects. MANDARINUNASP1 and UNASP4 are different. Despite the relatively large Δ log likelihoods (and negative weights) of these constraints, the phonological model predicts that numerically *fewer* syllables with the relevant onset type will bear the relevant tone than the standard syllables model does. While these reductions do bring the predicted rates closer to the observed rates, MANDARINUNASP1 and UNASP4 do not seem to represent straightforward phonological effects whereby, for example, loanword syllables with unaspirated onsets in Mandarin are more likely to bear high tone.

Finally, the stress- and position-related constraints STRESSED1, STRESSEDNON-FINAL1 and FINAL4 remain in the model after pruning, and appear to improve predicted rates of tonal assignment, but they have smaller weights and much smaller Δ log likelihoods than the voicing-related constraints.

3.4 Discussion

The MaxEnt analysis demonstrates that lexical tone distributions and standard syllables come close to explaining all tonal adaptation in the corpus; these are the chief determinants of tone in integrated loanwords. Additionally, however, there is a correlation between English voicing and tonal assignment. Syllables whose onsets adapt English voiceless consonants are more likely to bear high tone, while syllables whose onsets adapt English voiced consonants are more likely to bear rising tone; moreover, rising tone assignment is even more likely when onsets adapt English voiced *obstruents*. This last finding confirms a claim made by Wu (2006) on the basis of a small number of syllables, but while for Wu this was a secondary influence of the English input on tonal adaptation, subordinate to stress-to-tone, in my study it is the single largest phonological effect. The retention of the constraints STRESSED1, STRESSEDNON-FINAL1 and FINAL4 suggests there may be some slight stress-to-tone patterning within the already more marginal phonological determinants of tone, but

the contribution of these constraints is dwarfed by that of the onset voicing constraints. The overall results corroborate and sharpen Glewwe's (2016) claim of a voicing effect in established loanwords, strengthening the case by controlling for non-phonological factors in tonal assignment.[11] The MaxEnt analysis affirms that English onset voicing, not stress, is the dominant phonological factor shaping tonal adaptation in the loanword corpus.

Disentangled from the much more influential effects of the lexicon and standardisation, the voicing effect is a comparatively minor tonal adaptation pattern in established loanwords, yet it is nevertheless present. One way of interpreting it is within the perceptual model of loanword adaptation (Peperkamp & Dupoux 2003, Peperkamp 2005), in which mappings are established on the basis of phonetic similarity. English onsets cause F0 perturbations at the beginnings of syllables, such that F0 is higher after voiceless obstruents and lower after voiced obstruents and perhaps sonorants (Hombert *et al*. 1979; see §2 for discussion of the effect of sonorants on F0).[12] Mandarin-speaking adapters perceive this phonetic variation, and select the tones that best match these differing pitches early in the syllable: high tone for syllables whose onsets correspond to English voiceless obstruents, and rising tone for syllables whose onsets correspond to English voiced obstruents and sonorants. It may seem unlikely that a mapping from relatively subtle F0 perturbations to tone should be favoured over a mapping from the more salient stress-driven intonation patterns, but F0 perturbations are consistent across contexts, while the intonation patterns aligned with stress are not.[13] The pitch differences after different types of onsets may thus be the more reliable phonetic pattern, and therefore the preferred basis of a perceptual mapping.

Another possible interpretation of the voicing effect is Hsieh & Kenstowicz's (2008) contrast enhancement account for Lhasa Tibetan, which assigns tones to the initial syllables of Mandarin and English loanwords based solely on the voicing of the syllable's Mandarin or English onset. Voiceless onsets trigger high tone assignment and voiced onsets low tone assignment. Rejecting a perceptual mapping account, Hsieh & Kenstowicz instead propose that this adaptation pattern is rooted in contrast enhancement. Since tones do not serve to contrast lexical items within the loanword vocabulary, they can be recruited to enhance another phonological contrast, namely voicing. This occurs in Mandarin loanwords, even though Mandarin voicing (or sonority) is preserved in Tibetan segmental adaptation. The assigned Tibetan tones are an additional, albeit redundant, cue to the voicing feature.

---

[11] The voicing effect found here is consistent with a minor convention identified by Chang (2020) in her dictionary corpus, but that convention constitutes an isolated pattern (whether principled or accidental) embedded within the set of standard syllables. In contrast, the present voicing effect emerged after standard syllables were controlled for.

[12] Treating English as an aspirating language rather than a true voice language does not harm this account, because F0 is systematically lowered after English voiced stops even if they are phonetically voiceless, with short-lag VOTs (Kingston & Diehl 1994).

[13] I thank an anonymous reviewer for highlighting this point.

As in Tibetan, tone is likely not serving a contrastive function in the Mandarin loanword vocabulary. Native content words are typically disyllabic, and over half of disyllabic words have no segmental homophones, meaning they are unambiguous even without their tones (Lin 2017). Loanwords, particularly proper names, are often longer, so they will have even fewer segmental homophones, and tone will matter even less in their identification. Under Hsieh & Kenstowicz's contrast enhancement account, the voicing effect in my corpus does not stem from the perception of F0 perturbations, but is instead a strategy to enhance the contrast between voiced and voiceless consonants. As in the Mandarin-to-Tibetan case, the assigned tones reinforce a segmental contrast already generally preserved in adaptation.

A final possibility is that the voicing effect replicates a tone split in the development of modern Mandarin, perhaps suggesting that historical effects of onset consonants on tone that are no longer active in the native lexicon are still active in tonal adaptation. In Mandarin, syllables that in Middle Chinese bore the level (*ping*) tone acquired high tone if they had voiceless obstruent onsets and rising tone if they had voiced obstruent or sonorant onsets (Chen 2000), a split closely matched by the corpus voicing effect. While Li (2017) favours this account, it is unclear how this historical tone split could be reproduced in synchronic tonal adaptation by adapters who are unaware of how the Mandarin tones developed from the Middle Chinese ones.

There is a compelling reason to favour the perceptual mapping interpretation over both contrast enhancement and a historical effect, namely the strength of VOICEDOBS2. In the corpus, voiced obstruents in particular caused more rising tone assignment, beyond the general association between voicedness and rising tone. According to the contrast enhancement account, high and rising tones enhance a binary contrast between voiceless and voiced; this does not predict a difference between voiced obstruents and voiced sonorants. Similarly, the historical explanation cannot account for why English voiced obstruents predict rising tone assignment even more strongly, since the level tone split did not distinguish between sonorant and voiced obstruent onsets. If the voicing effect is instead rooted in perception, the fact that voiced obstruents depress F0 to a greater degree and/or more reliably than sonorants can neatly explain the importance of VOICEDOBS2 over and above VOICED2.

The fit between the observed and predicted tones in the final, phonological MaxEnt model ($r^2 = 0.967$) shows that this model performs very well, accounting for the vast majority of tonal assignment in the corpus. The remaining variation may be partly due to character avoidance (Chang 2020), whereby adapters try not to use characters with undesirable meanings (Li 2017) or characters for function morphemes that might cause speakers to hear unintended compounds or phrases embedded within a loanword (Chang 2020).[14] It would be difficult to implement a further

[14] Chang (2020) excluded some function morphemes in calculating lexical tone probabilities, but she later identifies additional function morphemes that were not excluded and that may have triggered character avoidance effects.

control for character avoidance in the corpus, however, because character avoidance is itself variable. For instance, Chang lists 可 *kě* 'can, may' as one character that is avoided in her corpora, but it appears several times in my corpus. Besides character avoidance, undetected semantic considerations or indirect borrowings (§ 3.1) may also underlie the remaining unexplained variation.

4　Loanword adaptation experiment

To complement the corpus study, I conducted an experiment eliciting native Mandarin speakers' online adaptations of English words. The experiment was modelled on Chang & Bradley (2012), and tested hypotheses developed from the corpus study and the previous literature. In the corpus, English voicedness (especially in obstruents) predicted more rising tone assignment, while English voicelessness predicted more high tone assignment. In terms of segmental effects, then, I hypothesised that adapted syllables whose onsets corresponded to English sonorants or voiced obstruents would prefer rising tone, while syllables whose onsets corresponded to English voiceless obstruents would prefer high tone. In light of the attention paid to stress-to-tone in the Chinese loanword adaptation literature, as well as the fact that adaptation patterns can differ between integrated loanwords and online adaptations, I also hypothesised that stressed English syllables would be more likely to receive high tone than unstressed English syllables.

4.1　Method

4.1.1　*Materials*.　The stimuli consisted of 80 disyllabic nonce English words, half trochees and half iambs; examples are given in (5).

(5) a. *Trochees*　　　　　　　b. *Iambs*

['mɑbə]	*maba*	[mə'bɑ]	*maba*
['bɑpi]	*bapi*	[bə'pʰi]	*bapi*
['pʰibə]	*piba*	[bə'mɑ]	*bama*
['fænʃəɹ]	*fanshi*	[pʰɪ'mi]	*pimi*
['pɪnjə]	*pinya*	[fən'jɑ]	*fenya*

They were constructed such that each syllable had an expected Mandarin segmental adaptation that could bear all four tones.[15] The onset consonants were sonorants (/m j/), phonemically voiced stops/affricates (/b g dʒ/), phonemically voiceless stops/affricates (/p tʃ/) or voiceless fricatives (/f s ʃ/). The phonemically voiced stops/affricates were further categorised as phonetically voiced when in the onset of the final (unstressed) syllable of a trochee, and as devoiced otherwise. The phonemically voiceless stops/affricates were further categorised as unaspirated

---

[15]　The expected segmental adaptations were based on my intuition.

when in the onset of the final (unstressed) syllable of a trochee, and as aspirated otherwise. Table II gives the distribution across onset types of the 160 syllables that made up the 80 disyllabic nonce words, with an example syllable of each type. The phonetic realisations of stimuli syllables are given in IPA, and the expected Mandarin segmental adaptations in Pinyin.

	initial		final		
	stressed	unstressed	stressed	unstressed	*total*
sonorant	[mɑ] *ma*	[mə] *ma*	[mɑ] *ma*	[mə] *ma*	28
voiced stop/affricate	—	—	—	[bə] *ba*	16
devoiced stop/affricate	[bɑ] *ba*	[bə] *ba*	[bɑ] *ba*	—	40
aspirated stop/affricate	[pʰi] *pi*	[pʰɪ] *pi*	[pʰi] *pi*	—	32
unaspirated stop/affricate	—	—	—	[pi] *pi*	16
voiceless fricative	[fæn] *fan*	[fən] *fen*	[fæn] *fan*	[fən] *fen*	28
total	40	40	40	40	160

Table II

Distribution of syllables composing the English nonce word stimuli.
IPA transcriptions give the phonetic realisations of sample syllables
in the English nonce words, and Pinyin indicates the expected
segmental adaptations of sample syllables in Mandarin.

For all but one of the English syllables, the expected segmental adaptation was the same whether the syllable was stressed, with a full vowel, or unstressed, with a reduced vowel (e.g. [ˈbɑ] and [bə] → *ba*). The syllable [fæn], which appeared as [fən] when unstressed, had the expected segmental adaptation *fan* in stressed position and *fen* in unstressed position.

The disyllabic nonce words were constructed by randomly concatenating possible initial and final syllables, with correction by hand to ensure that no item consisted of two syllables with the same expected segmental adaptation or corresponded to a real English word. Because the stimuli set was designed to include an equal number of trochees and iambs and some syllable types are positionally restricted, it was not possible for each onset type to be equally represented.

The stimuli were recorded by a phonetically trained male native speaker of American English, who was naive as to the purpose of the study. The recordings were produced in a soundproof booth using a Shure SM-10A head-mounted microphone plugged into an XAudioBox. The recordings were carried out using the program PCQuirerX, with a sampling rate of 22,050 Hz. Each nonce word was produced in the frame 'This is the word X. X.' Thus each word was uttered twice, first in sentence-final

position and then in isolation. The stimuli were embedded in a carrier sentence to encourage participants to consider them as words of English. The isolated token was inspired by Chang & Bradley's (2012) study, in which each nonce word was pronounced as if it were a sentence in itself. Both tokens were final in their intonational phrase. The intonation pattern for trochees was H* L-L%, with the H* pitch accent on the initial syllable. The intonation pattern for iambs varied between L+H* L-L% and H* L-L% (though phonetically these were very similar), with the pitch accent on the final syllable. Word-initial voiced stops and affricates were not prevoiced, and exhibited short-lag VOTs, but F0 was significantly lower after both voiced stops and affricates ($t(146.3) = -2.657$, $p = 0.009$) and sonorants ($t(142.9) = -4.000$, $p < 0.001$) than after voiceless stops and affricates.[16]

4.1.2 *Subjects.*   The participants were fifteen UCLA students (six male; nine female), ranging in age from 18 to 24 (median 18.5). All subjects were native speakers of Mandarin Chinese who had grown up in mainland China. The age at which they began learning English ranged from 5 to 12 (median 7). All subjects had been living in the United States for two years or less (median 6 months).

4.1.3 *Procedure.*   The experiment was implemented in E-Prime 2.0 and conducted in a soundproof booth. Subjects wore a Plantronics Audio 400 DSP headset. They were told that they would hear made-up English words, and that their task was to say and then write how they would most naturally adapt these English words into Mandarin. After a practice phase, subjects proceeded at their own pace through a single block of 80 trials corresponding to the 80 disyllabic nonce stimuli. The 80 trials were randomly ordered for each subject. Subjects could take a break after 40 trials. In each trial, subjects heard the nonce word in its frame. They were then asked to say out loud how they would adapt the nonce word into Mandarin. This oral response was recorded. When they had given their response, subjects were asked to type the adaptation they had just spoken in Pinyin, using numbers to indicate tone. This written response was also recorded. The experiment lasted 20 to 30 minutes.

4.1.4 *Data processing.*   Analyses were carried out on the individual syllables of the oral responses. Each subject's oral responses were independently transcribed by two native Mandarin speakers. The transcribers did not have access to the English stimuli, and were unaware of the experimental hypotheses. The average rate across subjects of inter-transcriber agreement was 91% for both segments and tones and 94% for tones only. While I analysed only the oral responses, the rate of agreement for

[16]   Mean F0, averaged across the first 11% of the vowel, was 122.5 Hz after voiceless stops and affricates, 113.8 Hz after voiced stops and affricates and 109.0 Hz after sonorants.

tone between subjects' transcribed oral responses and their own written responses was 95%.

I excluded eleven syllables for which the two transcriptions did not match and a third independent transcription could not confirm one of the two transcriptions as correct. I also excluded 50 syllables that were subject to third tone sandhi. There were 38 syllables that were unintelligible or not recorded because subjects began their oral responses early. Another 20 syllables were excluded because subjects perceived the nonce word stimulus as a real English word and translated it. The two stimuli that occasionally provoked this response were [dʒə'pɪn] (perceived as *Japan*) and [ʃəɹ'gɑ] (perceived as *cigar*). Finally, I excluded 27 tokens of *ce*, *fo*, *fou*, *me*, *ri* and *se*, because these syllables can only bear one tone in Mandarin. In total, I excluded 146 syllables, leaving 2254 analysable tokens.

4.2 Standard syllables in the experiment

As in the loanword corpus, there was evidence for the use of standard syllables in the experiment. I first identified oral responses that matched the pronunciation of the character prescribed for the given English stimuli syllable in the Xinhua News Agency's English-to-Mandarin transcription chart. For stressed syllables, I looked up the syllable in the chart. If an oral response matched the pronunciation of the character prescribed for that syllable, I coded it as using the standard syllable (e.g. 巴 *bā* is prescribed for /bɑː/, so any oral response *bā* for the stimulus ['bɑ] was considered standard). For unstressed syllables, I categorised more possible characters as standard, to account for the possibility of orthographic influence or misperception. While subjects never saw written forms of the stimuli, their knowledge of English may have caused them to make inferences about how the words they were hearing were spelled. Additionally, unstressed [ɪ] and [ə] are relatively confusable, and voiceless unaspirated onsets may have been misperceived as voiced. Consequently, I looked up characters prescribed for possible spellings or misperceptions of stimuli syllables, and coded oral responses as standard if they used the pronunciations of any of these characters. The stimulus syllable [ʃəɹ] does not have an adaptation in the chart, because [əɹ] is not one of the listed rhymes. Subjects usually adapted this syllable as *shi* or *she*. The chart includes standard characters with the pronunciations *shí*, *shě* and *shè*, so I coded such oral responses for [ʃəɹ] as standard.

In all, I identified 827 syllables (37% of oral responses) as corresponding to standard adaptations of the English stimuli syllables. Another 230 adapted syllables corresponded to *some* standard syllable in the Xinhua News Agency chart, even if it was not the standard syllable for the given English stimulus. The remaining 1197 syllables (53%) did not belong to the set of standard syllables.

While standard syllables were less ubiquitous in the experiment than in the corpus, they were nevertheless prevalent. Moreover, there may be a

greater need to control for standard syllables in the experiment, because the range of English syllables was so narrow, increasing the likelihood of spurious phonological effects. For example, ignoring vowel reduction, there were only four English syllables with voiceless stop or affricate onsets ([ʧʰɑ], [pʰi], [ʧʰi] and [pʰɪn]), and the standard adaptations of all four bear rising tone (*chá*, *pí*, *qí* and *píng*). In the following section, I first examine the non-standard syllables, in order to gain an initial idea of what phonological effects are at play in the experiment. I then include the standard syllables in the MaxEnt analysis, for the same reason as in the corpus analysis: while they reflect official standards, they may also participate in phonological patterning, and this information would be lost if they were excluded.

4.3 Initial results and MaxEnt analysis of the experiment

I first take a general look at tonal assignment in the 1197 non-standard syllables in the experiment. Figure 1 shows the proportion of syllables adapted with each of the five tones by stress and position in English. The graph is inevitably noisy, since there is no accounting for the lexical tone probabilities of the individual segmental syllables. Nevertheless, the figure suggests certain patterns. Stressed initial syllables overwhelmingly prefer high tone (0.71 *vs.* <0.13 for any other tone), while stressed final syllables strongly prefer falling tone (0.47 *vs.* <0.20 for any other tone). This looks like stress-to-tone: high tone captures the H* pitch accent aligned with an initial stressed syllable, while falling tone captures the H* L-L% contour realised on a final stressed syllable. Thus it appears plausible that English intonation, here operationalised as stress and position, influenced tonal adaptation in the experiment.
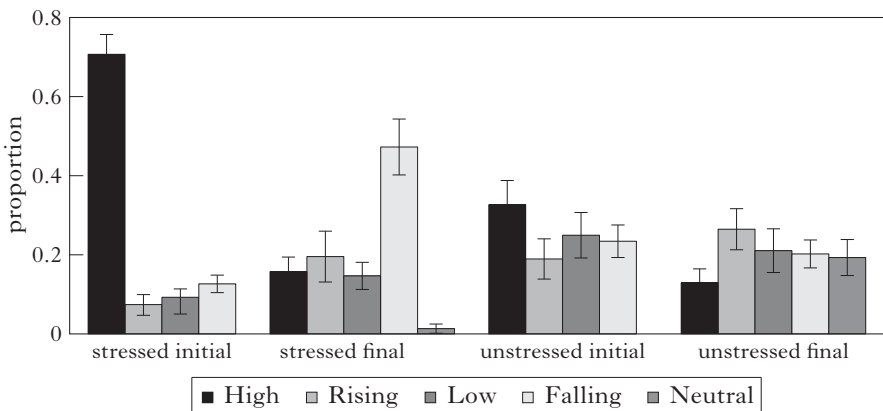

Figure 1

Tonal assignment by stress and position in non-standard syllables, averaged across participants.

The MaxEnt analysis of the experiment was identical to that of the corpus, except for minor differences related to the design of the experiment. From the 2254 oral responses, including standard syllables, I excluded the 72 syllables produced with neutral tone. Based on Fig. 1, neutral tone is used almost exclusively for unstressed final syllables. Adaptation with neutral tone is appropriate for unstressed final syllables, because Mandarin syllables with neutral tone are always unstressed and word-final (Duanmu 2007). It seems unlikely that any additional phonological factors (e.g. segmental properties) would favour or disfavour adaptation with neutral tone, since it is not a lexical tone with a fixed pitch level or contour, but rather varies in realisation, depending on the preceding tone. Since the corpus data consisted of written Mandarin forms, it contained no identifiable instances of neutral tone, so excluding neutral tone syllables in the experiment makes the two studies more comparable. I therefore conducted the analysis on 2182 oral responses. The English inputs were again abstract units distinguishing stress (stressed *vs.* unstressed), position (initial *vs.* final) and four onset types: sonorant, voiced stop/affricate, voiceless stop/affricate and voiceless fricative. Sub-tableaux were again organised by Mandarin segmental syllable.

In the first step of the analysis, I established the lexical model, which predicted what tones subjects' oral responses would have had if they had assigned tones based purely on the lexical tone probabilities of the segmental syllables they had chosen. The fit (r^2) between the observed counts in the study and the counts predicted by the lexical model was 0.574. Just as native tone frequencies could not fully account for tonal adaptation in the corpus, so they cannot fully explain which tones subjects used in the experiment.

In the second step of the analysis, I added standard syllable constraints to the MaxEnt model and determined their weights. In the standard syllables model, the fit between the observed and predicted counts improved ($r^2 = 0.722$). This shows that the tones of standard syllables did influence subjects' tonal adaptations, but unlike in the corpus, it is not the case that lexical effects and standardisation can explain nearly all tonal assignment.

The third step of the analysis investigated the additional role of phonological determinants of tone. I used the same 76 phonological constraints as in the corpus MaxEnt analysis (replacing NON-FINAL with INITIAL wherever it occurred, since all non-final syllables were initial in the experiment). After determining their weights, I removed 13 constraints that were assigned weights of zero. I then pruned the phonological model, using the same procedure as for the corpus. The final pruned model contained 28 phonological constraints. In this model, the fit between the observed and predicted tone counts improved further ($r^2 = 0.837$), and the 28 phonological constraints collectively contributed significantly to the model ($\chi^2(28) = 572.592$, $p < 0.001$). Table III lists the 28 phonological constraints with their weights, in descending order of Δ log likelihood.

constraint	weight	Δ log likelihood	predicted rate		observed rate
			standard syllables model	phonological model	
Unasp2	−1.084	59.789	25.6%	29.9%	29.5%
Initial1	−0.823	47.695	28.4%	46.6%	45.9%
StressedInitial1	−1.112	46.635	30.9%	63.6%	62.2%
StressedFinal4	−0.997	45.552	23.2%	32.7%	32.4%
Unasp4	−1.003	36.667	23.6%	17.7%	16.1%
VoicedObs1	−0.622	30.839	46.3%	48.1%	50.0%
Unstressed3	−0.768	30.527	15.8%	23.5%	23.9%
Obs1	−0.444	30.371	34.8%	39.5%	39.9%
Stressed1	−0.582	25.073	32.9%	46.9%	46.2%
Voiced3	−0.480	15.370	16.5%	18.3%	19.2%
Tone2	−0.292	13.737	28.8%	27.8%	27.9%
VoicedObs3	−0.617	11.685	7.0%	10.5%	11.3%
Asp4	−0.599	9.500	14.7%	14.0%	15.4%
MandarinAsp2	−0.411	9.087	52.1%	48.4%	48.2%
Voiced1	−0.237	7.040	34.5%	38.3%	39.4%
UnstressedFinal3	−0.640	6.465	17.1%	25.2%	25.0%
Stressed2	−0.227	6.167	28.1%	21.7%	22.4%
Obs3	−0.329	6.117	12.1%	13.4%	13.7%
Voiceless4	−0.343	6.097	21.3%	18.4%	18.7%
MandarinUnasp2	−0.354	5.298	18.1%	19.3%	19.5%
Asp3	−0.879	4.625	10.9%	12.2%	11.1%
Initial2	−0.179	3.764	29.2%	23.4%	24.0%
Unstressed2	−0.267	3.744	29.7%	34.4%	33.8%
Tone3	−0.198	3.737	16.2%	17.0%	17.4%
Son2	−0.317	3.413	26.1%	26.6%	26.6%
Final2	−0.217	2.941	28.4%	32.2%	31.7%
MandarinUnasp4	−0.439	2.857	26.4%	20.2%	17.9%
Son4	−0.395	2.266	27.4%	20.6%	20.1%

Table III
Phonological constraints in the experiment MaxEnt analysis.

Ignoring Unasp2, which I discuss below, the top three constraints are Initial1, StressedInitial1 and StressedFinal4, suggesting that English stress and position were the strongest phonological determinants of tone. Specifically, the initial syllables in English disyllables were overall more likely to receive high tone. On top of this effect, there was a particular preference for adapting *stressed* initial syllables with high tone.

Finally, stressed final syllables were more likely to be adapted with falling tone. Relative to the standard syllables model, the addition of the 28 phonological constraints does in fact increase the predicted counts of initial syllables with high tone, stressed initial syllables with high tone and stressed final syllables with falling tone, bringing these counts closer to the observed counts. Two other relatively important constraints, UNSTRESSED3 and STRESSED1, also behave straightforwardly in the model: relative to the standard syllables model, the phonological model predicts more unstressed syllables to have low tone and more stressed syllables to have high tone, in both cases achieving a closer match to the observed tones.

Among the top tiers of phonological constraints, there are also several that refer to onset properties. The importance of UNASP4 seems to lie in complex and opaque interactions with the rest of the model (cf. MANDARINUNASP1 in the corpus MaxEnt analysis), because the predicted rate of falling tone assignment in adapted syllables whose onsets correspond to English unaspirated onsets is not greater in the phonological model than in the standard syllables model. Rather, it is in fact smaller, coming closer to the observed rate. Thus UNASP4 does not correspond to a clear phonological effect in the way that the stress and position constraints above do.

The case of UNASP2 is different. The addition of the phonological constraints slightly increases the predicted rate of rising tone assignment in syllables whose onsets adapt unaspirated stops and affricates, and this increase leads to a better fit to the observed rate. In examining the model more closely, though, I discovered that what seemed to be driving the large negative weight and large Δ log likelihood of UNASP2 were standard syllables with rising tone whose associated standard syllable constraints had received quite low or zero weights in the standard syllables model. Additionally, many of these syllables' onsets corresponded to *voiceless* unaspirated stops and affricates, which only occurred in unstressed final position. To give the most dramatic example, which is representative of the wider phenomenon, the standard syllables model predicted that only 19 tokens of *pi* adapting English (unstressed final) [pi] would receive rising tone, while the phonological model, with UNASP2, predicted 29 rising tone tokens. The actual number was 30. Although *pí* is a standard syllable, the constraint PI2 had the 39th smallest weight out of the 45 standard syllable constraints that received any weight. The incidence of *pí* in stressed and/or initial syllables, where stronger stress-to-tone effects may have reduced standard syllable use, may not have justified assigning PI2 a larger weight. Consequently, the standard syllables model underestimated standard *pí* specifically in unstressed final syllables (the only English syllables with voiceless unaspirated onsets), which led to UNASP2 having to pick up the slack in the phonological model. In short, I contend that the importance of UNASP2 does not represent a genuine phonological effect, but is rather an artefact of how English aspiration is confounded with stress and position, the fact that standard syllables may have been used

more where stress-to-tone effects were weaker, and the tones of the standard syllables for the particular stimuli in the experiment. There is evidence that the importance of VOICEDOBS1 and OBS1, two constraints with Δ log likelihoods similar to those of UNSTRESSED3 and STRESSED1, can be explained in the same way.

4.4 Comparing the corpus and experiment analyses

The phonological models of the corpus (Table I) and the experiment (Table III) reveal sharp differences in the phonological effects that are active in each set of adaptations. The three constraints that contributed most significantly to the corpus model, VOICEDOBS2, VOICED2 and VOICELESS1, were not retained in the experiment model after pruning (i.e. they did not contribute significantly at the 0.05 level), indicating that these onset voicing effects played no meaningful role in participants' online adaptations. Meanwhile, of the top three stress- and position-related constraints in the experiment model, INITIAL1, STRESSEDINITIAL1 and STRESSEDFINAL4, only STRESSEDINITIAL1 (= STRESSEDNON-FINAL1) remained in the corpus model after pruning, where its contribution was less significant ($\chi^2(1) = 5.830$, $p = 0.016$ *vs.* $\chi^2(1) = 93.270$, $p < 0.001$ in the experiment model). Similarly, UNSTRESSED3 was not retained in the corpus model, while STRESSED1's contribution was less significant than in the experiment model ($\chi^2(1) = 7.522$, $p = 0.006$ *vs.* $\chi^2(1) = 50.146$, $p < 0.001$). The stress-to-tone effects that prevailed in online adaptations (see below) were thus either absent or weaker (especially compared to other phonological factors) in integrated loanwords.

4.5 Discussion

As in established loanwords, lexical tone frequencies and the influence of standard syllables were major determinants of tone in the experiment, though their combined role was not as overwhelmingly dominant as in the corpus. In considering the phonological determinants of tone, I return to the hypotheses put forward at the beginning of §4. First, the voicing effect observed in the corpus did not emerge in the experiment. Adapted syllables whose onsets corresponded to English sonorants or voiced obstruents were not more likely to receive rising tone, and adapted syllables whose onsets corresponded to English voiceless obstruents were not more likely to receive high tone. The prediction that stressed syllables would be more likely to receive high tone was borne out, but the picture is more nuanced, with both stress and position implicated. Being initial in an English disyllable increased the likelihood of high tone assignment. Additionally, there were interactions of stress and position, such that high tone was more likely specifically in stressed initial syllables, and falling tone more likely specifically in stressed final syllables. Two lesser effects were an increased likelihood of low tone on all unstressed syllables and of high tone on all stressed syllables.

Together, these patterns suggest that English intonation was the principal phonological determinant of tone in online adaptations. This can be accounted for by the perceptual mapping model. In choosing tones, subjects sought to make their adaptations phonetically similar to the stimuli by reproducing the English stress-aligned pitch contours as best they could. The pitch accent on stressed initial syllables translated to high tone, while the pitch accent and subsequent fall on stressed final syllables translated to falling tone. The effects of INITIAL1, UNSTRESSED3 and STRESSED1 can be interpreted as grosser generalisations, perhaps reflecting a preference for simpler mappings. High tone could adapt the pitch accent of a stressed initial syllable or the unspecified pitch of a pretonic unstressed initial syllable. Low tone could adapt the relatively low pitch of unstressed syllables, whether they are initial, and therefore less prominent than the following stressed syllable, or final, realising the L-L% of the English intonation pattern. High tone could also be used generally for all stressed syllables, which bear the pitch accent, even if stressed final syllables also contain a pitch fall. There is almost no evidence that subtler pitch perturbations following different types of onsets shaped subjects' tonal assignments. Instead, it was the overall intonation of the English forms that mattered.

In the corpus MaxEnt analysis, adding the phonological constraints increased the fit (r^2) between predicted and observed tones only from 0.952 to 0.967, but in the experiment analysis, the fit increased from 0.722 to 0.837. This suggests that phonological effects played a bigger role in the online adaptations than in established loanwords. At the same time, the final model accounted for more tonal assignment in the corpus than in the experiment, where more variation remains unexplained. As in the corpus, some of the remaining variation may be due to subjects avoiding characters (even though they did not have to give responses in Chinese characters) or avoiding the production of real Mandarin words. Other possible sources of variation in the experiment include subjects' individual tonal adaptation preferences and a potential self-priming effect whereby subjects tend to adapt the next word with the same tones used on the previous word. Because the input data to the MaxEnt analysis were not coded for subject or trial number, any such patterning would have been inaccessible to the model, impeding its ability to explain all tonal assignment.

5 General discussion

I have presented a corpus study and an online adaptation experiment investigating Mandarin tonal adaptation. My corpus was comparatively large by the standards of the literature. I examined stress in integrated loanwords much more comprehensively than previous studies; this work is also the first to analyse unstressed syllables in online adaptations. To tackle the notorious intractability of Mandarin tonal adaptation, I used

MaxEnt modelling. By building lexical tone distributions directly into the model and then successively modelling the effects of standard syllable use and phonological factors on top of the lexicon, much more of the variation in tonal assignment could be accounted for than in previous studies. Moreover, I demonstrated that, after lexical statistics and the effects of standardisation were controlled for, phonological properties of the English input still played a role in tonal adaptation.

In both the corpus and the experiment, tonal assignment was greatly influenced by lexical tone probabilities, consistent with Chang's (2020) claim that native tone frequencies are the primary mechanism underlying tonal adaptation. Nevertheless, lexical factors still only explained around 60% of the variation in assigned tones in each study. The use of standard syllables also explained much of the tonal assignment in both studies, particularly in the corpus. While Chang (2020) identifies certain prescriptive conventions, she stops short of recognising the full extent of standardisation in Mandarin loanword adaptation. Other studies (Mar & Park 2012, Zheng & Durvasula 2016) have also acknowledged the role of standardisation, without being able to fully quantify or control for it.

In established loanwords, the lexicon and standard syllable use accounted for nearly all tonal assignment. Nevertheless, phonological effects played a small but significant role: loanword syllables whose onsets corresponded to English voiceless consonants were more likely to have high tone, while syllables whose onsets corresponded to English voiced consonants, especially voiced obstruents, were more likely to have rising tone. While previous studies have noted onset effects in Mandarin tonal adaptation, I attribute the effect specifically to the English form, identify English voicing as the crucial factor, especially in obstruents, and argue that the effect stems from mapping English F0 perturbations to tone.

Phonological effects played a larger role in the online adaptations. Mandarin speakers chose tones that best reproduced the intonation patterns of the English stimuli. Previous studies have uncovered variable and sometimes contradictory stress-to-tone effects in Mandarin tonal adaptation; my results are most similar to Zheng & Durvasula's (2016) experimental finding that English–Mandarin bilinguals preferred to adapt initial stressed syllables with high tone.

The phonological effects in the corpus and the experiment results can both be understood within the perceptual model, yet they are quite different. The degree of variation in the source forms may explain why the perceptual mapping to tone is from F0 perturbations in the corpus, but from English intonation in the experiment. Participants in the experiment heard only two tokens of each stimulus, each produced with the same canonical intonation pattern. Since the stress-aligned pitch contours are more salient than the onset-induced F0 perturbations, it makes sense that participants would seek to imitate the former in assigning tones. In natural speech, on the other hand, the same English word can occur with different intonation patterns, depending on context, and so, for

example, the pitch accent aligned with the stressed syllable may not always be H*. Meanwhile, F0 perturbations are consistent: regardless of intonation pattern, F0 will always be higher after a voiceless stop than after a voiced stop. In the context in which the established loanwords were borrowed, then, adapters may have based their perceptual mapping on the more stable pitch property, namely F0 perturbations.

The differences between the two studies demonstrate how the perceptual model can yield different adaptation patterns for the same language pair, depending on the precise situation in which adaptation is occurring. These differences point to the importance of the wider borrowing context, and, combined with other recent studies, suggest that Mandarin tonal adaptation may be in flux. In terms of integrated loanwords, Mandarin tonal adaptation has always differed strikingly from Cantonese tonal adaptation, which largely follows stress-to-tone principles. This divergence can plausibly be attributed to dissimilar borrowing contexts. The Cantonese studies examine English loanwords in Hong Kong Cantonese. Hong Kong experienced over a century of British colonial rule, during which English and Cantonese were in direct and sustained contact (Li 1999, Hao 2009). This close contact seems to have given rise to a mapping from English intonation to Cantonese tone, despite the variability of intonation patterns across contexts. Additionally, adaptations were apparently not constrained by the writing system, since some loanword syllables violate native phonotactics and tonotactics (Yip 2006). In contrast, mainland China, where Mandarin is spoken, was never ruled by English-speaking colonisers in the sustained way that Hong Kong was (Miao 2005). The loanwords in my corpus were mostly borrowed in the 19th century in a massive wave of translating English written works, or in the late 20th century, as China opened up to the West (Miao 2005, Dong 2012). In the former period, adaptations may have been based largely on English written forms. In the latter period, Mandarin speakers were increasingly in contact with English, but not in a colonial context. Additionally, all loanwords in my corpus were written in Chinese characters. All this could help explain why the vast majority of tonal adaptation in the corpus was not phonological.

The borrowing context of the experiment was one of forced close contact, and participants did not have to write their adaptations with characters. Though artificial, the situation was closer to the borrowing context of Cantonese, and indeed, there was more phonologically driven tonal adaptation, which exhibited stress-to-tone patterns. Unlike the results of the experiment, though, Cantonese stress-to-tone cannot be attributed to the fixed intonation patterns of the English forms. If stress-to-tone patterns can emerge even from variable input, why do the minor phonological effects in Mandarin established loanwords not also mainly reflect stress-to-tone? I conjecture that the reason is again a lack of sufficiently close and sustained contact. What little perceptual mapping occurred in the integrated loanwords was from consistent English F0 perturbations, not from canonical stress-driven intonation patterns extracted from variable

data. It may be that this extraction, and therefore dominant stress-to-tone adaptation, requires close contact in the borrowing context.

Some evidence in support of this idea comes from recent studies on Mandarin tonal adaptation in 'lettered words', i.e. words or word parts written in the Latin alphabet even in written Chinese (e.g. *X guāng* 'X-ray', *NASA*, *wi-fi*) (Ding & Li 2020, Wang 2020). That lettered words can remain in the Latin alphabet is itself indicative of a closer contact situation, since Mandarin speakers must know the English letter names or be able to read English words. Additionally, while this is still an emergent strand of research, there are preliminary indications that lettered words exhibit more robust stress-to-tone patterns than older established loanwords like those in my corpus. That is, their tonal assignment patterns are more similar to those in my experiment. While it is too early to be certain, closer contact between English and Mandarin, fostered by the spread of English-language media, may be going hand in hand with more stress-to-tone adaptation, and the next wave of borrowings may show stronger influences of stress-to-tone.

From the present investigation, it appears that Mandarin adapters rely in no small part on lexical tone distributions when assigning tones to loanwords. They are further influenced by the prescriptive standards they have absorbed, and also apply perceptual mappings from pitch properties of the English input to tone. Future work could develop a tonal adaptation grammar that incorporates these factors and is shaped by the specific borrowing context.

Because it can handle variable outputs and a many-level response variable like lexical tone, the MaxEnt approach holds promise for other cases of tonal adaptation which exhibit considerable variation and/or in which multiple determinants of tone are at play. For instance, in Thai (Kenstowicz & Suchato 2006) and White Hmong (Golston & Yang 2001), multiple properties of the English input affect tonal assignment. MaxEnt modelling could shed new light on such cases by clarifying the relative influences of these properties. It could also be used to incorporate fine-grained lexical statistics directly into the analysis in cases where lexical tone distributions are hypothesised to influence tonal adaptation.

REFERENCES

Chang, Zhuting (2020). *Tonal adaptation of loanwords in Mandarin: phonology and beyond*. PhD dissertation, City University of New York.

Chang, Zhuting & Dianne Bradley (2012). Tonal assignment for English loanwords in Mandarin: an experimental approach. Paper presented at the CUNY Conference on the Segment. Abstract available at https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.432.4721&rep=rep1&type=pdf.

Chen, Matthew Y. (2000). *Tone sandhi: patterns across Chinese dialects*. Cambridge: Cambridge University Press.

Cheng, Chin-chuan (1968). English stresses and Chinese tones in Chinese sentences. *Phonetica* **18**. 77–88.

Collins, Beverley & Inger M. Mees (1996). *The phonetics of English and Dutch*. Leiden: Brill.

Da, Jun (2004). A corpus-based study of character and bigram frequencies in Chinese e-texts and its implications for Chinese language instruction. In Pu Zhang, Tianwei Xie & Juan Xu (eds.) *Proceedings of the 4th International Conference on New Technologies in Teaching and Learning Chinese*. Beijing: Tsinghua University Press. 501–511.

Davis, Stuart, Natsuko Tsujimura & Jung-yueh Tu (2012). Toward a taxonomy of loanword prosody. *Catalan Journal of Linguistics* **11**. 13–39.

de Jong, Kenneth & Mi-hui Cho (2012). Loanword phonology and perceptual mapping: comparing two corpora of Korean contact with English. *Lg* **88**. 341–368.

Ding, Hongwei & Yiling Li (2020). Tonal adaptation of disyllabic letter-character pattern in Mandarin alphabetical words. In *Proceedings of the 10th International Conference on Speech Prosody*. 494–498. http://doi.org/10.21437/SpeechProsody.2020-101.

Dong, Xiaoli (2012). *What borrowing buys us: a study of Mandarin Chinese loanword phonology*. PhD dissertation, University of Utrecht.

Duanmu, San (2007). *The phonology of Standard Chinese*. 2nd edn. Oxford: Oxford University Press.

Glewwe, Eleanor (2016). Tonal adaptation in English loanwords in Mandarin: a corpus study. *CLS* **51**. 197–211.

Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a Maximum Entropy model. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.) *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Stockholm: Stockholm University. 111–120.

Golston, Chris & Phong Yang (2001). White Hmong loanword phonology. In Caroline Féry, Antony Dubach Green & Ruben van de Vijver (eds.) *Proceedings of HILP 5*. Potsdam: University of Potsdam. 40–57.

Hao, Yen-Chen (2009). Tonal adaptation of English loanwords in Cantonese. In Yun Xiao (ed.) *Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21)*. Smithfield, RI: Bryant University. 42–54.

Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *LI* **39**. 379–440.

Hayes, Bruce, Colin Wilson & Ben George (2009). Maxent grammar tool. Software package. Available (May 2021) at http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/.

Hombert, Jean-Marie (1976). The effect of aspiration on the fundamental frequency of the following vowel. *BLS* **2**. 212–219.

Hombert, Jean-Marie, John J. Ohala & William G. Ewan (1979). Phonetic explanations for the development of tones. *Lg* **55**. 37–58.

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

Hsieh, Feng-fan & Michael Kenstowicz (2008). Phonetic knowledge in tonal adaptation: Mandarin and English loanwords in Lhasa Tibetan. *Journal of East Asian Linguistics* **17**. 279–297.

Kang, Yoonjung (2010). Tutorial overview: suprasegmental adaptation in loanwords. *Lingua* **120**. 2295–2310.

Kenstowicz, Michael (2006). Tone loans: the adaptation of English loanwords into Yoruba. In John Mugane, John P. Hutchison & Dee A. Worman (eds.) *Selected Proceedings of the 35th Annual Conference on African Linguistics*. Somerville, MA.: Cascadilla. 136–146.

Kenstowicz, Michael & Atiwong Suchato (2006). Issues in loanword adaptation: a case study from Thai. *Lingua* **116**. 921–949.

Kingston, John (2011). Tonogenesis. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.) *The Blackwell companion to phonology*. Malden, MA: Wiley-Blackwell. 2304–2333.

Kingston, John & Randy L. Diehl (1994). Phonetic knowledge. *Lg* **70**. 419–454.

Kubozono, Haruo (2006). Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua* **116**. 1140–1170.

Leben, W. R. (1996). Tonal feet and the adaptation of English borrowings into Hausa. *Studies in African Linguistics* **25**. 139–154.

Lee, Wai-Sum & Eric Zee (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association* **33**. 109–112.

Li, David C. S. (1999). Linguistic convergence: impact of English on Hong Kong Cantonese. *Asian Englishes* **2**. 5–36.

Li, Jian (2017). Tonal adaptation of English loanwords in contemporary Mandarin. *Journal of Chinese Linguistics* **45**. 423–450.

Li, Kening (2005). Pitch perturbation of sonorants and a three-way tonal split in Chinese. *BLS* **31**. 443–452.

Lin, Isabelle (2017). *Tone sequences in lexical processing of Beijing Mandarin*. MA thesis, University of California, Los Angeles.

Maddieson, Ian (1997). Phonetic universals. In William J. Hardcastle & John Laver (eds.) *The handbook of phonetic sciences*. Oxford & Cambridge, MA: Blackwell. 619–639.

Mar, Li-Ya & Hanyong Park (2012). Tonal adaptation of English loanwords in Mandarin: the role of perception and factors of characters. Poster presented at the 162nd Meeting of the Acoustical Society of America, Kansas City, MO. Available (May 2021) at https://www.researchgate.net/publication/230860079.

MDBG (2014). *Chinese–English Dictionary*. www.mdbg.net.

Miao, Ruiqin (2005). *Loanword adaptation in Mandarin Chinese: perceptual, phonological and sociolinguistic factors*. PhD dissertation, Stony Brook University.

Ohde, Ralph N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *JASA* **75**. 224–230.

Osatananda, Varisa (1996). An analysis of tonal assignment on Japanese loanwords in Thai. In *Pan-Asiatic Linguistics: Proceedings of the 4th International Symposium on Language and Linguistics*. Nakhon Pathom: Institute of Language and Culture for Rural Development, Mahidol University at Salaya. 198–210.

Peperkamp, Sharon (2005). A psycholinguistic theory of loanword adaptations. *BLS* **30**. 341–352.

Peperkamp, Sharon & Emmanuel Dupoux (2003). Reinterpreting loanword adaptations: the role of perception. In M. J. Solé, D. Recasens & J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Causal Productions. 367–370.

Shinohara, Shigeko (2000). Default accentuation and foot structure in Japanese: evidence from Japanese adaptations of French words. *Journal of East Asian Linguistics* **9**. 55–96.

Silverman, Daniel (1992). Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* **9**. 289–328.

Wang, Zhi-Hao (2020). Tonal adaptation of Chinese lettered words. Ms, Peking University.

Wu, Hsiao-hung Iris (2006). Stress to tone: a study of tone loans in Mandarin Chinese. *MIT Working Papers in Linguistics* **52**. 227–253.

Xinhua News Agency Proper Names and Translation Service (1993). *Names of the world's peoples: a comprehensive dictionary of names in Roman-Chinese*. Beijing: Zhongguo Duiwai Fanyi Chuban Gongsi.

Xu, Ching X. & Yi Xu (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association* **33**. 165–181.

Yip, Moira (2006). The symbiosis between perception and grammar in loanword phonology. *Lingua* **116**. 950–975.

Zheng, Mingzhe & Karthik Durvasula (2016). English loanwords in Mandarin Chinese: a perception experiment approach. In Hongyin Tao, Yu-Hui Lee, Danjie Su, Keiko Tsurumi, Wei Wang & Ying Yang (eds.) *Proceedings of the 27th North American Conference on Chinese Linguistics*. Vol. 2. Los Angeles: University of California, Los Angeles. 462–480.