The replacement of Beijing Mandarin neutral tones in AABB reduplications

Sixing Cui and Michael J. Kenstowicz
Central China Normal University and MIT

1 Introduction

Earlier research (Cui 2012, 2021) suggests that when Mandarin disyllabic adjectives of the form AB such as [dàfang] 大方 'generous' are reduplicated to AABB 大大方方, the B syllable of the base is placed in a stressed position. Mandarin neutral tones (T0) are banned from stressed positions. What tone is assigned to the T0 syllable when an AB base like *dàfang* is reduplicated? We present the results of a study that poses this question for seven Beijing Mandarin native speakers. Our results indicate that the original tone associated with the T0 character/morpheme helps the speaker choose among the four possible substitutes for the T0: specifically, when the original tone of the B morpheme is T3 [214] or T4 [51] the speaker is significantly more likely to substitute T3 or T4, while original T2 [35] is more likely to change to T1 [55]. We interpret this difference as a conflict between faithfulness to an original (remote) base and a markedness preference for a high tone in a stressed syllable. We then investigate whether there are traces of the original tone in the acoustic properties of the neutral-tone base form that could help the speaker recover the original tone.

2 Word stress and the neutral tone

The existence of stress and its phonological and phonetic correlates is one of the most contentious issues in Chinese phonology. For example, Duanmu (1999, 2014), Hsieh (2021), and Ma (2021) argue for the presence of lexical and phrasal stress in Mandarin, while Zhang (2017), Zhou (2018, 2021), and Li & Wu (2022) dispute this view. There are several reasons why Chinese stress is so problematic. Due to its monosyllabic morpheme structure and minimal affixation, Chinese lacks long words that can give rise to the rhythmic alternations between weak and strong syllables that underlie the stress contours of many languages such as English: cf. *Màssachúsetts*, *Càlifórnia*, *Àpalàchicóla*, etc. Also, compounding rather than affixation is the principal word-building mechanism in Chinese. Hence the language lacks stress shifts like English *átom* ≈ *atóm-ic* and *rígid* ≈ *rigíd-ity* that make the contrast between stressed and unstressed syllables particularly vivid. Finally, in Mandarin a bimoraic requirement is imposed on the syllable rhyme (Duanmu 2008, Wu & Kenstowicz 2015) that minimizes potential differences in syllable weight that often correlate with word-level stress distinctions.

The one place where the phonological and phonetic evidence converge to motivate a stress contrast at the word level in Mandarin Chinese is the *qīngshēng*

轻声 neutral tone (NT, T0). The Mandarin neutral-tone syllable is comparable to an English unstressed syllable: it is shorter in duration, the vowel space is more compact, and it is often the site of segmental contractions and coalescence (Zhang et al. 2023). There is a large literature on the f0 properties of the neutral tone with debate on whether it is underspecified and acquires its f0 contour from the preceding syllable by carry-over coarticulation or alternatively has a dedicated pitch target in the mid-low region of the tonal space: Chao (1948, 1964), Li (2003), Lee & Zee (2008), Zhang (2021), among others. A recent typology (Chang 2024) divides the Mandarin neutral tone into three subcategories based on its grammatical source as a suffix (e.g. perfective -*le*), a cliticized particle (e.g. genitive #*de*), or the reduction of a full syllable in the second position of a lexical compound (e.g. *gāo-xing* 'happy'). The reduction to a neutral tone creates phonological opacity for the Tone-3 sandhi process in various grammatical constructions such as the reduplication seen in *jiáng-jiang* 'speak for a little bit' from underlying /jiǎng-jiǎng/ (Gao et al. 2019). The neutral tone is transcribed as a distinct phonological category in many dictionaries and is recognized in lists of normative pronunciation. Finally, the neutral tone is banned from the initial syllable of the word as well as from a stressed syllable.

3 Research question

In her study of phrase-level stress in Chinese, Cui (2012, 2021) finds that adjectives such as *ānshēng* 安生'peaceful' have a SW-WS stress rhythm (S = strong, W = weak) when reduplicated: 安安生生 *ānān-shēngshēng* 'quite peaceful'. If this premise is accepted then it implies that when an AB base adjective with a neutral tone like *bièniu* 'awkward' is input to this grammatical construction to create AABB, a conflict will arise between the phonology and the grammar. Faithfulness to the phonology wants to reduplicate the neutral tone to yield *bièbiè-niuniu*. But the grammar of this construction assigns a prosodic template requiring the B syllable to be stressed. Since a neutral-tone syllable cannot be stressed, the grammar is faced with a contradiction. How is this conflict resolved? As far as we know, this question has not been investigated before and is our research question in this paper.

Pretheoretically, there is a range of possible resolutions of the conflict. AABB reduplication might be blocked for AB bases with a final T0, creating a paradigm gap. This is not found. As far as we know, speakers do not hesitate to reduplicate AB neutral-tone bases any more than they do AB bases with a full tone. If the grammar reduplicates an AB adjective regardless of the prosodic status of its second syllable, then is the ban on stressed neutral tones suspended in the AABB construction? Here again the answer is "No". The data indicate that the neutral-tone ban is still enforced. If the ban on stressed neutral tones is respected then the T0 of the neutral-tone syllable must change to a full tone in the AABB construction. What full tone is chosen as a substitute for the neutral tone?

Theoretically, there is a range of possible neutral-tone substitutions one might expect to occur. The most obvious would be a default tone that is optimal in the stressed position (markedness based). Alternatively, the speaker/grammar might take recourse to frequency matching and choose the most common tone, either in general or in this particular construction. Still another possibility is faithfulness to a more remote form: the B lexical item (character or morpheme) may have a full tone alternant in some other realization (e.g. the citation form of the bare neutral-tone syllable) or in a nonreduced compound. This would be analogous to a faithfulness effect tied to or cued by the writing system, as in English *J[ə]pán* ≈ *Jàpan-ése*, *N[ə]pál* ≈ *Nèpal-ése*, and *M[ə]lán* ≈ *Mìlan-ése* where affixation of stressed *-ése* creates a clash with the final stress of the base noun that is repaired by shifting the stem stress onto the reduced schwa-like vowel of the first syllable necessitating the substitution of a full vowel [æ], [ɛ], [ɪ]. Finally, the choice of which full tone to substitute for the neutral tone might be guided by phonetic properties of the T0 syllable of the AB base itself, possibly containing 'traces' of the underlying tone of the B-syllable morpheme/character. The question we are asking here resembles cases in which the base-form for a morphophonological alternation such as the citation form is also a phonological neutralization site. A well-studied example is the neutralization of the coronal obstruent codas in the nominal inflection of Korean (Albright 2002, Kang 2003, Jun 2009, Ito 2009, Do et al. 2014, and more recently Jo 2024). The coda of the Korean syllable is the site of the merger of the five phonemic obstruents /s, t, t^h, c, c^h/ to an unreleased stop [t̚]: cf. *nat̚* ≈ *nas-il* 'sickle', *nat^h-il* 'piece', *nac-il* 'day', *nac^h-il* 'face'. It turns out that among the five possible alternants that can emerge before a case suffix like the accusative *-il* c. 70% have t ≈ s, reflecting a lexical imbalance inherited from Middle Korean (Ito 2010). The t̚ -> s rule attracts infrequent stems and serves as a model for loanword inflection (Ito & Kenstowicz 2017): *t^{h}ik^{h}et, t^{h}ik^{h}es-il* 'ticket'.

Among the four phonemic tones of Mandarin (T1 [55], T2 [35], T3 [21] ≈ [214], and T4 [51] represented as H, LH, L, and HL in Yip's (2002) OT analysis), T1 is the most likely default candidate. It accords with de Lacy's (2002) finding of a cross-linguistic preference for the alignment of high tones and stressed syllables. Also Zhang (2021) reports that when a Mandarin word with neutral tone is placed in a syntactically focused position the neutral tone is changed to T1. With respect to lexical frequency, our query of the Modern Chinese Dictionary found 3,867 disyllabic AB adjectives. Some 244 are listed with a neutral tone in the second syllable. Among these c. 150 can reduplicate as AABB in the BCC (Beijing Language & Culture University 现汉和等级划分轻声词) corpus. The B-syllable original tones in the AABB construction are distributed as follows: T1 (N=35), T2 (N=45), T3 (N=16), and T4 (N=51). The recent study of Yang et al. (2022) finds the following frequencies among the four lexical tones in general: T1 (105,000), T2 (96,000), T3 (129,000), and T4 (228,000). If the Mandarin speaker

was consulting the frequencies of their grammar's lexicon, we would expect T4 to predominate as the best choice for a prosodically licit stand-in for the neutral tone.

4 Study design

Since the reduplicated AABB form of an AB adjective is not normally recorded in dictionaries, we solicited our data from a pool of seven Beijing Mandarin native speakers. Their gender and age profiles run as follows: F1 (25 yrs), F2 (40), F3 (30), F4 (22), M1 (22), M2 (50), M3 (25). Our corpus consists of 40 items drawn from the BCC list of 150 AB adjectives with a neutral tone. The 40 items were divided into groups of ten based on the original tone associated with the neutral-tone B syllable that appears in either a citation bare form or in another compound where the original tone was transcribed. In addition, we selected 16 AB adjectives with a B-syllable full tone, also divided into four groups based on the tone of the B syllable with four examples for each of the four basic tones. See the appendix for the list of 56 items comprising our AB word-list. The entire list was randomized and presented to our speakers in Chinese characters inside a frame sentence *wǒ shuō X zhège cí* 'I say X (with) this character', which the speaker was asked to read. Each AB prompt was followed by the AABB reduplicated form in the same frame sentence, also presented in Chinese characters. For example, for the adjective 客气 *kèqi* 'polite', the speaker recorded 我说客气这个词 [wǒ shuō kèqi zhège cí] followed by 我说客客气气这个词 [wǒ shuō kèkeqīqī zhège cí]. The recordings were made in the MIT Linguistics Dept Speech Lab with a Shure SM10A Unidirectional Head-Worn Dynamic Microphone and a USB Pre-2 Preamp at a sampling rate of 44.1 kHz and quantizing resolution of 16 bits.

Each of the authors independently transcribed the tones assigned by our speakers to the AB and AABB forms. The second author's transcriptions were checked by a native Mandarin-speaking linguistics graduate student familiar with the Mandarin tones and their transcription. The two data files differ slightly in that the first had some repetitions that were absent from the second file. In general, the data trends were the same across both files.

5 Results

The tables below show the pooled counts of the tones assigned to the B syllables of the reduplicated adjectives as a function of the original tone associated with the B syllable of the AB base. Table 1 indicates the tone assigned in the reduplicated form when the B syllable is a full tone. The first column in the table shows the tone of the B syllable in the base and the top row indicates the four possible full tones that could show up in the reduplicated form. The cells at the intersection indicate the counts for each condition. From this table it is clear that when the B syllable of the AB base has a full tone then it is almost always faithfully reduplicated. Faithful copy occurs uniformly with T2 and T3. For T4 there are three cases where T1 is found instead of the expected T4. Finally, when the base contains T1 in the B position there are three anomalous responses: two with T2

and one with T4. In general then, when the B syllable of the AB base has a full tone, it is faithfully copied in the AABB reduplicated form. We should mention here that when the B tone is T3 then the T3-sandhi rule regularly applies to the reduplicated BB string yielding a T2T3 output. Here and in what follows we count these cases as faithful copy.

Table 1: B = full tone

A**B** / AA**BB**	T1	T2	T3	T4
T1	27	2		1
T2		26		
T3			28	
T4	3			25

Table 2 below shows the behavior of neutral-tone AB bases where the T0 of the base must be replaced by a full tone in the reduplicated form. The table format is the same as Table 1 with the first column indicating the original tone associated with the B-syllable character in the base in some related form and the top row showing the four possible full-tone substitutes. In the case of Table 2 a quite different pattern of correspondences is found.

Table 2: B = neutral tone (first transcription file)

A**B** / AA**BB**	T1	T2	T3	T4		T1	T2	T3	T4
T1	120				120	1.0			
T2	92	28			120	.76	.24		
T3	48		72		120	.40		.60	
T4	53			67	120	.45			.55
totals	313	28	72	67	480				

The following observations can be made concerning Table 2. First, the neutral tone of the AB base is replaced by T1 or the original tone, never by an ‘anomalous’ tone. Second, when the underlying character-tone is T2 then the T0 of the AB base is changed to T1 92 times in AABB and appears as T2 28 times. But when the underlying character-tone is T3 or T4 then there is a bias to substitute in the reduplicated form the tone that is associated with the character: 72 times for original T3 and 67 times for original T4. But T1 substitutions are found as well: 48 times for original T3 and 53 times for original T4. The second half of the table shows the proportions of the responses among the four possible choices.

Table 3 displays the pattern of neutral-tone substitutions in the file transcribed by the second author. Overall, there is greater faithfulness to the original tone associated with the character compared to Table 2. But there is still a bias to replace original T2 with T1 (.70) while original T3 and T4 prefer to be faithful to the original tone (.70 and .62, respectively).

Table 3: B = neutral tone (second transcription file)

original/AABB	T1	T2	T3	T4		T1	T2	T3	T4
T1	74				74	1.0			
T2	53	22	1		76	.70	.29	.01	
T3	23		53		76	.30		.70	
T4	28		1	47	76	.37		.01	.62
totals					300				

6 Summary

Stepping back from the details, we can summarize our findings as follows. When the AB base contains a full tone in B it is faithfully copied in AABB. When the AB base contains a neutral tone in B it is always replaced by a full tone. In the overwhelming majority of cases the full-tone substitute is either T1 [55] or the tone associated with the B-syllable character. This variation can be interpreted as a conflict between markedness (high tone in a stressed syllable) vs. faithfulness to a remote form of the base cued by the character. The choice between markedness vs. faithfulness correlates to some extent with the underlying/original tone of the neutral-tone syllable: original T2 (LH) favors the change to T1 while original T3 (L) and T4 (HL) are biased against change and in favor of faithfulness. The faithfulness hierarchy to the original-B tone response T3 > T4 > T2 does not track the type frequencies mentioned above, where T4 is the most frequent B tone in the AABB construction as well as more generally in Mandarin. Hence, we conclude that our speakers are not matching the frequency of the lexicon for the AB > AABB adjectival construction; something else is going on.

7 Statistical tests

We ran several binomial mixed-effects logistic regressions in R to gauge the statistical significance of our results—in particular the difference between original tone T2 vs. T3 and T4. In these tests the dependent variable is Faithful with the values “No”/0 if the transcribed tone differs from the original tone and “Yes”/1 if the transcribed tone matches the original tone. The “No” response is the baseline in the models. The independent predictor variable is the original tone of the B character. We excluded cases where the original tone was T1 since a T1 response is ambiguous between a markedness-driven change and faithfulness. Speaker and word were our random intercepts.

Model 1= glmer(faithful ~ originalB + (1|speaker) + (1|word), data = data,
family = "binomial") AIC = 191.5 BIC = 207.5

	Estimate	Std Error	zvalue	Pr(>|z|)
(Intercept)	-2.6395	0.8813	-2.995	0.00275 **
original-BT3	3.7282	1.1763	3.169	0.00153 **
original-BT4	2.9475	1.0843	2.718	0.00656 **

Model 1 suggests that when the predictor variable is changed from baseline original-B = T2 to original-B = T3 then there is a significant increase in the faithfulness = 'Yes' response; a similar but weaker effect holds for T4.

Model 2 below added the token log frequency of the AB base as a predicting factor. It did not significantly improve model fit (ch-sq = 2.5, p = 0.11).

Model 2= glmer(faithful ~ originalB + log-base-freq + (1|speaker) + (1|word), data = data, family = "binomial") AIC = 190.9 BIC = 210.1

	Estimate	Std Error	zvalue	Pr(>|z|)
(Intercept)	-0.4642	1.5335	-0.303	0.762108
original-BT3	4.4273	1.2881	3.437	0.000588 ***
original-BT4	3.6181	1.1724	3.086	0.002029 **
log-base-freq	-0.8067	0.5109	-1.579	0.114326

Model 3 below added the cumulative log frequency of the B-syllable character to the original model as a predictor. This time the additional factor was significant. The effect is negative, suggesting that the lower the position of the B character is in the frequency rank (and hence the higher its numerical value) the more likely the substitution of T1 for T0 is. Stated differently, the more frequent the character, the more likely it is to inhibit the markedness-driven substitution of T1 for T0 and instead accord with faithfulness.

Model 3= glmer(faithful ~ originalB + log-char-cum-freq + (1|speaker) + (1|word), data = data, family = "binomial") AIC = 188.7 BIC = 207.9

	Estimate	Std Error	zvalue	Pr(>|z|)
(Intercept)	5.901	4.085	1.444	0.14
originalBT3	4.780	1.330	3.595	0.0003 ***
originalBT4	2.897	1.044	2.775	0.0055 **
log-char-cum-fr	-4.967	2.402	-2.068	0.0386 *

8 Discussion

Several factors might play a role in explaining the greater propensity for original T2 neutral-tone syllables to change to T1 vis à vis original T3 and T4. First, both T1[55] and T2 [35] are realized primarily in the upper region of the pitch space (Xu 1997). Second, T1 and T2 derive from the same Middle Chinese tone class (*píng* 平 level). Third, according to the confusion matrix in Figure 1 below, Mandarin T2 (Rising) is more likely to be confused with T1 (High) than with T3 (Dipping) or T4 (Falling). These considerations suggest that the change from T2 to T1 is smaller and so may be 'less noticeable' (Steriade 1997, 2009) in the Beijing Mandarin speaker's internal grammar than the change of T3 or T4 to T1.

Figure 1: tonal confusion matrix for Mandarin (Wu et al. 2013)

(%)	High	Rising	Dipping	Falling
High	70.4	10.9	5.4	13.3
Rising	21.7	59.6	8.6	10.2
Dipping	16.9	25.9	44.2	13.0
Falling	18.4	7.2	11.5	63.0

Another possibility is that the neutral tone of the AB base may contain vestiges of the original tone that could help guide the speaker in the choice of the stressed-syllable replacement for the AABB reduplicated form. To pursue this point, we excerpted the neutral-tone syllables of the AB bases in our corpus and analyzed the data with the help of the Prosody-Pro Praat script (Xu 2013). The data were categorized by the original tone of the character forming the second syllable of the AB base. Table 4 below shows the mean values obtained for various f0 measures along with duration and intensity. The neutral-tone syllables associated with original T2 stand out from the other neutral tones for all factors except intensity. In particular, the neutral-tone syllables associated with original T2 have higher f0 values and shorter duration compared to original T3 and T4.

Table 4: neutral-tone syllables (prosodic measures)

originalB	f0 (Hz)	maxf0	minf0	excursion	duration(ms)	intensity (dB)
T1	186	218	155	5.69	145	63.9
T2	**217**	**272**	**180**	**6.65**	**110**	62.0
T3	199	228	166	5.35	129	60.7
T4	205	219	164	4.83	116	62.9
grd_mean	196	229	162	5.50	125	62.73

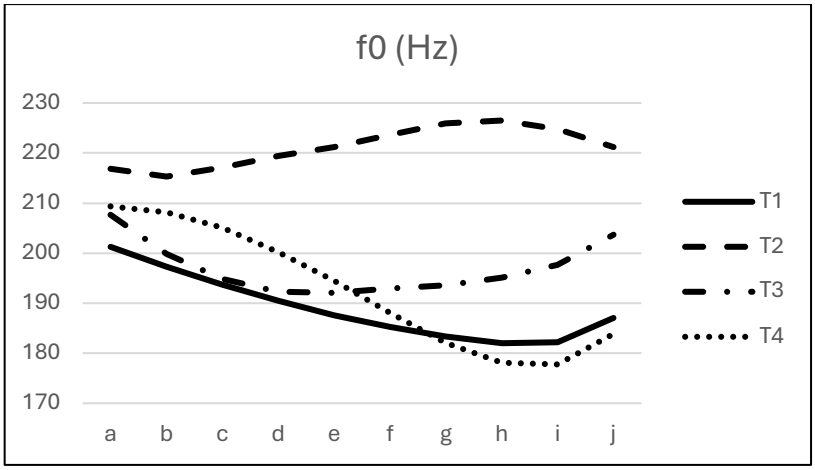
We ran a series of mixed-effects linear regressions (sum coding) on each of these factors: e.g. lmer(meanf0~originalB+(1|speaker)+(1|word), data=d). The t-values obtained for each test are shown in Table 5 below. They suggest that the neutral-tone syllables associated with original T2 have significantly higher f0 maxima and shorter durations compared to the overall average. The higher f0 of original T2 neutral tones might make the change to T1 [55] appear smaller in our speakers' linguistic consciousness compared to original T3 or T4 and hence could explain the differences in faithfulness between original T2 neutral-tone syllables vs. T3 and T4 seen in Tables 2 and 3 above.

Table 5: logistic regressions for neutral-tone syllables (t-values)

original B	f0 (ms)	maxf0	minf0	excursion	duration	intensity
T1	-1.01	-0.98	-1.17	-0.07	**4.16**	1.77
T2	1.43	**2.57**	1.87	1.13	**-3.0**	-0.65
T3	-0.09	-0.16	0.18	-0.09	0.54	-2.08

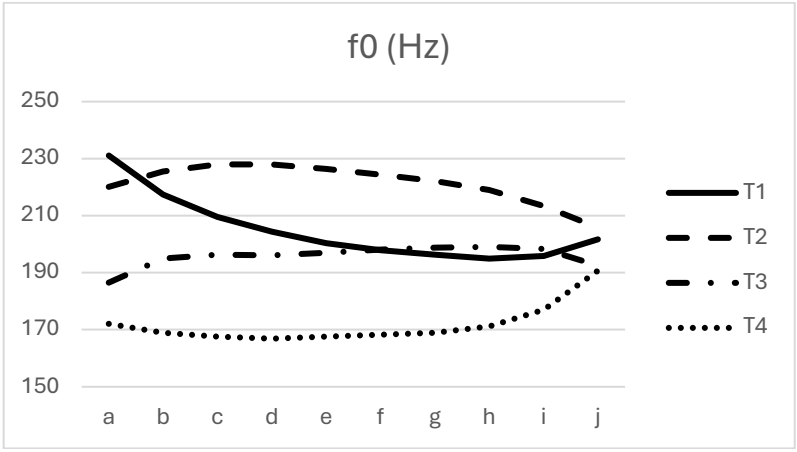
Higher f0 for the neutral-tone syllables associated with original T2 is also suggested by the time-normalized plots of the neutral-tone syllables shown in Figure 2 below. The horizontal axis of this plot shows the ten measurement points across the syllable rhyme and the vertical axis indicates the averaged f0 for each such point pooled across all speakers.

Figure 2: ten f0 timepoints of AB neutral-tone syllables coded by original-B tone



The plot suggests that neutral-tone syllables associated with original T1 and T4 fall to the low-tone target reported in earlier literature while original T2 and T3 reflect the rising and dipping trajectories of their underlying source. For comparison, Figure 3 below plots the same neutral-tone syllables but this time as a function of the tonal category of the preceding A syllable in the compound. In agreement with the earlier literature, this plot suggests that the relative height of the T0 is being carried over from the latter portion of the preceding syllable with high vs. low f0 values after the rising vs. falling trajectories of T2 and T4.

Figure 3: ten f0 timepoints of the neutral-tone syllables coded by A-syllable tone



In summary, one possible interpretation of the data is that the placement of the neutral tone in the upper vs. lower region of the pitch space is based on carry-over coarticulation from the preceding syllable while the rising, falling, or dipping trajectory reflects a trace of the original tone underlying the T0. We should point out that most of the earlier studies of the Mandarin neutral tone have looked at the behavior of inflectional particles where the original tone is distantly removed if recoverable at all. For our study the neutral-tones are drawn from lexical compounds where the original B-syllable tone is presumably more vivid in the speaker's consciousness. However, we should caution that there is a confound in the original-B T2 neutral-tone items in our corpus. The majority have a voiceless-consonant onset and high-vowel nucleus that could cue higher f0 and shorter vowel duration: shí 实, tú 涂, hú 糊. These are precisely the properties that singled out the T0 from original B that substituted T1 seen in Table 5 above. Hence the 'bias' for original-B T2 could be, at least partially, an effect of the item tested. But we also find substitutions of T1 for neutral tones derived from original T3 and T4 when the onset is not voiceless (e.g. 别扭 *bièniu* 'awkward'; see appendix 2) or the nucleus is not a high vowel (e.g. 地道 *dìdao* 'authentic'). So the change to T1 cannot be attributed entirely to an item effect or an onset effect.

9 Summary and Conclusions

This paper reports and discusses the results of a study in which seven Beijing Mandarin speakers were recorded pronouncing an AB adjective base along with its corresponding AABB reduplicated form. The B syllable of the base varied between a full tone and a neutral tone. Given that the B syllable of the base occupies a stressed position (weak or strong) in the reduplication construction, a neutral tone in the base must be changed to a full tone in the reduplicated form in order to conform to the undominated constraint against stressed neutral tones. The choice of which full tone to substitute exhibits a conflict between faithfulness to the original tone and change to the optimal T1 [55] high tone under stress. Our major finding is that a T0 associated with original T2 is biased to change in pursuit of markedness while original T3 and T4 favor faithfulness. We speculated that the change of T2 to T1 may be less noticeable in the speaker's internal grammar than the change of T4 to T1 or T3 to T1.

Further research into this topic might consider controlling the effect of the writing system by presenting the AB base stimuli aurally. A more balanced word-list that controls for the onset consonant of the neutral-tone syllable is a needed improvement along with a larger number of subjects and more 'objective' transcriptions of their responses. Finally, a more sophisticated modeling of the neutral tone with machine learning, Max-Ent grammars, or Gradient Symbolic Representations should be considered.

References

Albright, A. 2002. The Identification of Bases in Morphological Paradigms. UCLA Ph.D. dissertation.

Chao, Y-R. 1948. *Mandarin Primer*. Harvard University Press.

Chao, Y-R. 1964. *Grammar of Spoken Chinese*. University of California Press.

Chuang, J. 2024. Mandarin neutral tones as unstressed tones: Grammaticalization, stress, and tone sandhi. Lingbuzz.

Cui, S. 2012. From the stress-type of ABAB and AABB and its syntactic form to research the Chinese function of prosodic morphology. *Language Teaching and Linguistic Studies* 5. 63-69.

Cui, S. 2021. A study of stress patterns and stylistic attributes of Chinese reduplication. *Language and Linguistics* 22:3. 377-409.

de Lacy, P. 2002. The interaction of tone and stress in Optimality Theory. *Phonology* 19:1. 1-32.

Do, Y., C. Ito & M. Kenstowicz. 2014. The base of Korean noun paradigms: evidence from tone. *Korean Linguistics* 16:2. 111-143.

Duanmu, S. 1999. Metrical structure and tone: Evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics* 8. 1-38.

Duanmu, S. 2008. *Chinese*. Cambridge University Press.

Duanmu, S. 2014. 重音理论及汉语重音现象 [Stress theory and stress phenomena in Chinese]. *Contemporary Linguistics* 16:3. 288-302.

Gao, F., S. Lyu & C-J Lin. 2019. Tonal Processing in Mandarin Reduplication: Morphological and Lexical Effects. HISPhonCog 2019: Hanyang International Symposium on Phonetics & Cognitive Sciences of Language. Hanyang University, Seoul.

Hsieh, F-f. 2021. 论汉语有词重音 [On the existence of word stress in Mandarin Chinese]. *Studies in Prosodic Grammar* 8. 27-58.

Ito, C. 2010. Analogy and lexical restructuring in the development of nominal stem inflection from Middle to Contemporary Korean. *Journal of East Asian Linguistics* 19. 357-383.

Ito, C. & M. Kenstowicz. 2017. Pitch accent in Korean. *Oxford Research Encyclopedia of Linguistics*, Mark Aronoff ed. DOI: 10.1093/acrefore/9780199384655.013.242.

Li, Z. 2003. The Phonetics and Phonology of Tone Mapping in a Constraint-Based Approach. MIT Ph.D. dissertation.

Li, Y., & Wu, Y. 2022. 迷失在逻辑中: 论“北京话是一个重音语言”之说 [Lost in logic: On Feng's (2016, 2021) claim that Beijing is a stress language]. *Contemporary Linguistics* 24:5. 764-780.

Jun, J. 2010. Stem-final obstruent variation in Korean. *Journal of East Asian Linguistics* 19. 137-179.

Kuo, J. 2023. Phonological Markedness Effects in Reanalysis. UCLA Ph.D.

dissertation.
Lee, W-S & E. Zee. 2008. Prosodic characteristics of the neutral tone in Beijing Mandarin. *Journal of Chinese Linguistics* 36. 1-29.
Ma, Q. 2021. 汉语的词重音及其音系性质 [Chinese word stress and its phonological properties]. *Studies in Prosodic Grammar* 8:2. 112-129.
Xu, Y. 1997.Contextual tonal variations in Mandarin. *Journal of Phonetics* 25. 61-83.
Xu, Yi. 2013. ProsodyPro — A tool for large-scale systematic prosody analysis. In Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France: 7-10.
Wu, J., S. Zahorian & H. Hu. 2013. Tone recognition for continuous accented Mandarin Chinese. Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988. [https://www.researchgate.net/profile/Hongbing-Hu]
Wu, F. & M. Kenstowicz. 2015. Duration reflexes of syllable structure in Mandarin. *Lingua* 164. 87-99.
Yang, T-H., S-J. Jin & Y-A. Lu. 2022. The effect of lexicality, frequency, and markedness on Mandarin tonal categorization. *Frontiers in Psychology* 13.
Yip, M. 2002. *Tone*. Cambridge University Press.
Zhang, H. 2017. *Syntax-phonology Interface: Argumentation from Tone Sandhi in Chinese Dialects*. Routledge.
Zhang, Y. 2021. Neutral Tone in Mandarin: Representation and Interaction with Utterance-level Prosody. Cambridge University Ph.D. dissertation.
Zhongpei Z., A. Lacheret-Dujour & F. Isel. 2023. Phonetic correlates of lexical neutral tone within Mandarin two-character compounds in four full tones contexts. *International Journal of Language and Linguistics* 11:4. 107-118.
Zhou, R. 2018. 争议与思考 60 年来汉语词重音研究述评 [A review of Mandarin word stress study over 60 years]. *Language Teaching and Linguistic Studies* 194:6. 102-112.
Zhou, R. 2021. 汉语词重音问题再论 [A further discussion on word stress in Mandarin]. *Studies in Prosodic Grammar* 8:2. 150-170.

Appendix-1: Test Items

number	AB base		gloss	tone	original-B	AABB
1	抠嗦	kōusuo	stingy	1+0	1	抠抠嗦嗦
2	勤谨	qínjin	diligent	2+0	3	勤勤谨谨
3	别扭	bièniu	awkward	4+0	3	别别扭扭
4	和气	héqi	kind	2+0	4	和和气气

5	公道	gōngdao	fair	1+0	4	公公道道
6	刺痒	cìyang	itchy	4+0	3	刺刺痒痒
7	憨实	hānshi	honest	1+0	2	憨憨实实
8	大方	dàfang	generous	4+0	1	大大方方
9	滑溜	huáliu	smooth	2+0	1	滑滑溜溜
10	规矩	guījù	disciplined	1+0	3	规规矩矩
11	光溜	huáliu	smooth	1+0	1	光光溜溜
12	厚实	hòushi	thick	4+0	2	厚厚实实
13	委屈	wěiqu	aggrieved	3+0	1	委委屈屈
14	热火	rèhuo	enthusiastic	4+0	3	热热火火
15	麻利	máli	swift	2+0	4	麻麻利利
16	胡涂	hútu	confused	2+0	2	胡胡涂涂
17	肥实	féishi	fat	2+0	2	肥肥实实
18	敦实	dūnshi	stocky	1+0	2	敦敦实实
19	清楚	qīngchu	clear	1+0	3	清清楚楚
20	脆生	cuìsheng	crispy	4+0	1	脆脆生生
21	富态	fùtai	plump	4+0	4	富富态态
22	白净	báijing	white	2+0	4	白白净净
23	欢实	huānshi	vivacious	1+0	2	欢欢实实
24	抠搜	kōusou	stingy	1+0	1	抠抠搜搜
25	红火	hónghuo	lively	2+0	3	红红火火
26	含糊	hánhu	ambiguous	2+0	2	含含糊糊
27	干巴	gānba	dry	1+0	1	干干巴巴
28	冷清	lěngqing	deserted	3+0	1	冷冷清清
29	利索	lìsuo	nimble	4+0	3	利利索索
30	地道	dìdao	authentic	4+0	4	地地道道
31	花哨	huāshao	fanciful	1+0	4	花花哨哨

32	对付	duìfu	makeshift	4+0	4	对对付付
33	客气	kèqi	polite	4+0	4	客客气气
34	马虎	mǎhu	careless	3+0	3	马马虎虎
35	粗拉	cūla	careless	1+0	1	粗粗拉拉
36	糊涂	hútu	confused	2+0	2	糊糊涂涂
37	迷糊	míhu	confused	2+0	2	迷迷糊糊
38	瓷实	císhi	fat	2+0	2	瓷瓷实实
39	筋道	jīndao	chewy	1+0	4	筋筋道道
40	宽敞	kuānchang	broad	1+0	3	宽宽敞敞
41	安生	ānshēng	peaceful	1+1		安安生生
42	蓬松	péngsōng	fluffy	2+1		蓬蓬松松
43	普通	pǔtōng	ordinary	3+1		普普通通
44	认真	rènzhēn	earnest	4+1		认认真真
45	缠绵	chánmián	affectionate	2+2		缠缠绵绵
46	公平	gōngpíng	impartial	1+2		公公平平
47	渺茫	miǎománg	vague	3+2		渺渺茫茫
48	太平	tàipíng	peaceful	4+2		太太平平
49	彻底	chèdǐ	thorough	4+3		彻彻底底
50	懒散	lǎnsǎn	lazy	3+3		懒懒散散
51	和美	héměi	harmonious	2+3		和和美美
52	纷扰	fēnrǎo	tumultuous	1+3		纷纷扰扰
53	白胖	báipàng	chubby	2+4		白白胖胖
54	方正	fāngzhèng	upright	1+4		方方正正
55	冷落	lěngluò	neglected	3+4		冷冷落落
56	浩荡	hàodàng	vast	4+4		浩浩荡荡

Appendix-2: base bièniu ‘awkward’ and reduplicated bièbie-niūniū (Praat waveforms and labeled spectrograms)