EXPLAINING MANDARIN T3 SANDHI WITHIN AN ARTICULATORY PHONOLOGY FRAMEWORK

Shan Luo
University of Victoria

1. Introduction

Tonal systems in Chinese are quite complex and intricate (Chen, 2000). Among various Mandarin tonal alternations, the most well-studied pattern is Tone 3 Sandhi (T3S): Tone 3 becomes Tone 2 when followed by another Tone 3. The rule is given in (1) and illustrated in (2), drawn from Duanmu (2007, p. 256).

(1) T3 ⟶ T2/__T3
(2) a. Word: ma3 yi3 ⟶ ma2 yi3 "ant"
b. Compound: mi3 jiu3 ⟶ mi2 jiu3 "rice wine"
c. Phrase: ni3 hao3 ⟶ ni2 hao2 "You good (how are you?)"

This well-known phenomenon applies whenever the two low tones (T3) are adjacent in a domain, where the domains are defined by "a complex mixture of syntactic, prosodic, and focus-related factors" (Kuo, Xu & Yip, 2007, p. 211). Besides T3S, T3 also undergoes tone reduction word-initially: it becomes a half T3 when followed by T1, T2 or T4. The rule responsible for this tone-reduction alternation is known as the Half T3 Sandhi (Half T3S), which is "optional" according to Chen (2000). The below example in (3) shows a complete T3 ([214[] becomes a half tone ([21]) when followed by a T1 ([55]).

(3) Half T3 Sandhi (Chen, 2000, p. 21)
xiao mao "little cat" /214.55/ ⟶ /21.55/

The T3 variations, particularly T3S, are of so much interest to researchers that a very extensive literature has been dedicated to this subject (see Duanmu, 2007; Lin, 1992, 1998; Zhang, 2002; Yip, 2002, among others). For example, Dell (2004) suggests that T3S is a case of neutralization or reduction due to the first syllable in a /T3.T3/ sequence lacking stress. However, Duanmu (2004) shows it is possible to put stress on either syllable in a sequence like */mai3.ma3/* ("buy horse"), and argues T3S applies regardless of which syllable has more stress. Cheng (1973), on the other hand, argues that T3S is a case of dissimilation: no identical tones are allowed in a sequence. Following Cheng's analysis, Yip (2002) asserts T3S is "clearly dissimilatory" (p. 181) and refers to it as an OCP (Obligatory Contour Principle) effect (Leben, 1973; McCarthy, 1979, 1986): a sequence of two identical tones is prohibited.

However, the well-adopted OCP/dissimilation theory alone cannot account for three other questions: a) why sequences of two identical tones other than T3 are allowed; in other words, /T1.T1/, /T2.T2/, /T4.T4/ are allowed; b)

Actes du congrès annuel de l'Association canadienne de linguistique 2013.
Proceedings of the 2013 annual conference of the Canadian Linguistic Association.

why the sandhi tone is T2; c) why the output is not */T3.T2/, even though such an output also satisfies OCP by having the second T3 undergo the change? The first question is also addressed in Duanmu (2007), which states "there is no good answer" and the reasons for T3S are left open (p. 255). In spite of the wide-spread discussion T3S has received in research, how to apply AP hypotheses in a T3S analysis has not yet been extensively examined.

To preview the arguments in this study, I propose three main points. Firstly, following previous phonetic observations I argue that Half T3S is not optional but rather obligatory. Secondly, T3 is different from other Mandarin tones in terms of articulatory features: T3 is the only tone that has laryngeal constriction. Thirdly, I propose four articulatory constraints that impose duration (Tonal-DUR), constriction (OCP-Cons) and perceptual ($MAX_{initial}$) constraints as well as reflect temporal gestural overlap (MAX_T) within the framework of AP. The interaction among these constraints accounts for the three previous unattended questions. Considering all Chinese dialects are subject to a disyllabic requirement by which a minimal expression should consist of two syllables (Duanmu, 1994), and that disyllabic words constitute the greater part of the Mandarin Chinese lexicon (Chao, 1968), the scope of this study will focus on disyllabic T3 variations. Also, because of the scope of this study, neither syntactic branching nor prosodic domains will be relevant as the change can cross any syntactic domain in disyllabic words as shown in (2).

2. Background

2.1. Introduction of Mandarin Tonology

As a tonal language, Mandarin[1] consists of four lexical tones, which can be described by pitch contour (level, rising, falling) and pitch register (high, low, mid) (Lin, 1992; Xu, 1994). They are Tone 1 (high level, H, [55]), Tone 2 (rising, R, [35]), Tone 3 (dipping contour, fall then rise, L, [213/4]), and Tone 4 (falling, F, [51]) (see Chao, 1964; Xu, 1994). The four lexical tones vary in duration and are generally ranked in the following order from the shortest to the longest duration when produced in isolation: T4, T1, T2, and T3 (Gao, 2008; Tseng, 1990; Xu, 1997). In connected speech, however, Wang and Wang (1993) and Howie (1976) argue that the duration of T3 is not significantly longer than other tones.

As introduced earlier, Chinese sandhi systems are quite intricate when the adjacent tones come into contact with each other in connected speech (Chen, 2000). Depending on the position of which syllable preserves the tone of the canonical form, Chinese tone sandhi systems are often classified as left-dominant or right-dominant. Left-dominant sandhi often involves rightward extension (e.g., Shanghai). Right-dominant sandhi languages, on the other hand, often involve insertion and neutralization of nonfinal tones (Lin, 1994, 2001; Zhang, 2007). Mandarin dialects have been argued to be typical cases of right-

[1] Mandarin in this article refers to both the standard language (Putonghua) and Beijing Mandarin (the dialect.)

dominant tone sandhi systems, where the initial syllable, often being unstressed and relatively short, is not capable of bearing a concave tone.

The reason for this analysis is mainly due to T3 in Mandarin undergoes reduction/change in word-initial position: Half T3S and T3S. For Half T3S, T3 ([214]) is reduced to a half T3 ([21]) word-initially (see (3)). However, it is a complete tone in word-final position (Chao, 1968; Yip, 1980; Lin, 1983; Hashimoto, 1987). For example, *jiu213 yue53* ("September") becomes *jiu21 yue53* but *bai35 jiu213* ("white wine") does not change. Another sandhi process supporting the statement that Mandarin is right-dominant is T3S: T3 becomes T2 when followed by another T3 (see (1) and (2)). In other words, T3 undergoes neutralization in non-final positions; it is the word-final tone that has an ability of bearing a complete concave tone (T3).

In this study, I will approach T3 variations in the point view of Articulatory Phonology (AP) and argue that Half T3S is not optional but rather obligatory both word-initially and word-finally. As a result, Mandarin is in fact not right-dominant (see Duanmu, 2004; Lin, 1994, 2001). This argument is based on two phonetic facts a) tonal coarticulation is of greater magnitude and duration progressively than regressively; b) T3 is always reduced to a half T3 regardless of its position in connected speech (Xu, 1997).

2.2. Introduction of AP

The framework of AP has been proposed as an account of how spoken language is structured (Browman & Goldstein, 1986, 1992, 1995; Byrd, 1996; Byrd & Saltzman, 2003). The basic units under AP are gestures, which are assumed to have intrinsic duration and to occur in space (Browman & Goldstein, 1992). Building on a version of the gesture model developed by Browman and Goldstein, Gafos (2002) proposes a framework accounting for the cross-linguistic differences in gestural coordination in consonant clusters. Differing from the traditional analysis of treating schwa insertion in Moroccan Colloquial Arabic (MCA) as the result of epenthesis, Gafos analyzes it as arising from gestural mistiming: consonantal gestures in clusters that are not sufficiently overlapped.

Specifically, in the scenario of coordination between two heterorganic consonants, they are produced with a low degree of overlap, and thus there is an acoustic release between them: schwa. Such a coordination relation is to align the C^1's oral gestures in the middle point with the onset of C^2's. Gafos names as CC-COORD. However, when the two consonants are identical, e.g., [$T^{ə}T$], "the gestures must be timed so that the onset of the second /t/ begins at some point late in the release phase of the first /t/" (Gafos, 2002, p.272). To solve the conflict between OCP, that is, no overlap of identical oral gestures, and CC-COORD, the two tip-blade gestures have to drift away and a schwa is inserted.

The central claim in Gafos's proposal is that any theory of phonology must include a notion of temporal coordination of gestures. In the present analysis, we will examine whether the central hypothesis of Gafos's model—gestural coordination between segments—is applicable to analyze Mandarin T3 variations.

2.3 Articulatory Features of Tone

In Gao's (2008, 2009) discussion of tonal alignment under AP, the tones in Mandarin can be considered as T gestures entering into a C-center typology with the C(onsonant) and V(owel) gestures. C and T gestures are coupled in-phase to the vowel but anti-phase to each other, which is similar to the C-center coupling relationship in consonant clusters discussed in Gafos (2002). Gao thus suggests tones may be treated as ***additional consonants*** (emphasis mine) at the level of gestural planning. However, Gao's model leaves aside an essential question in AP: what articulators underlie the T gestures. The importance of identifying tone articulatory features is because T3 is distinct from other tones: it is the only tone that has constriction.

On the one hand, larynx movement has long been considered responsible for pitch variation in producing tone (Honda, Hirai, Masaki, & Shimada, 1999; Moisik, Lin, & Esling, 2010). The recent laryngoscopy study conducted by Moisik et al (2010) and an X-ray investigation by Wang and Kong (2010) further support the hypothesis that there is a rather stable relation between larynx height and pitch. However, the relation is not this straightforward with regards to T3, which is a concave tone in citation form ([214]) that starts with a low pitch and moves lower then rises. The data in Wang and Kong shows that the beginning F0 of T3 falls while the larynx moves down, as expected. What is interesting is that at the next stage the F0 of T3 remains low, but the larynx moves ***down or up***. Finally, the F0 rises while the larynx moves up. Overall, they conclude "the larynx moves up or down as F0 rises and falls" (p. 337). The study of Moisik et al (2010) again corroborates the relationship between larynx height and pitch target. For example, larynx raising occurs for all three participants during T2 and larynx lowering during T4. More importantly, the parameter of laryngeal constriction is found to play a role in the production of Mandarin T3, which means two distinct patterns of laryngeal activity during low pitch targets (i.e., T3): "either larynx lowering with an unconstructed laryngeal setting or larynx raising with laryngeal constriction" (p.6).

It is important to point out this dual usage of larynx in T3 because the strong constriction degree only occurs in the production of low tonemes ([21]) when the larynx moves up/down, the larynx moves up in the latter stage when producing the rising contour of T3 ([14]). That being said, the dual usage of larynx is only exhibited in the low tonal domain. Recall that there is an arbitrary stage found in Wang and Kong's (2010) study where the participant either raises the larynx or lowers it in the middle of producing T3. This could be accounted for by assuming that the participant is constricting the larynx. In other words, while T1, T2, and T4 only employ CL (constriction location), larynx height, T3 does employ CL and CD (constriction degree) at the same time. In fact, the constriction gesture accompanying a low tone is exhibited not only in Mandarin. In Bai, a Sino-Tibetan language spoken in Yunnan Province, the contrast between "ji 21" (*bracelet*) and "ji 31" (*soil*), besides the pitch register (pitch height), is that the former (the lower tone) is accompanied by laryngeal constriction. To sum up the phonetic observation in studies we have reviewed in this section, it is the larynx as the articulator combining height and constriction that is responsible for MC tone production. The act of producing Mandarin tones thus can be decomposed into articulatory gestures as follows:

(4) Laryngeal Gestures of Tone

T1: [larynx high + level]
T2: [larynx raising]
T3: [larynx constriction + larynx raising] or [larynx lowering]
T4: [larynx lowering]

Note that T3 has two alternatives: people either raise the larynx accompanying with constriction or lower the larynx. One might argue that simply lowering larynx would suffice, which seems having more articulatory ease. However, as we explained, constriction is generally associated with low tonemes cross-linguistically as is most clearly seen of T3 in the study of Moisik et al (2010). In this study, we propose that T3 is more likely to employ the first case, constriction+raising, instead of lowering larynx, to achieve the low tonal values of [21] in T3.

3. Articulatory constraints

3.1. Tonal-DUR

Regarding Half T3S and T3S, the most important question is what triggers them. In this section, I propose it is likely time driven (also see Lin, 1992). I formalize this claim as a constraint called Tonal-DUR.

It has been argued that there is no free-articulatory-association: articulatory mechanisms in producing pitch (F0) in fact impose many limitations on the way F0 ***can be*** produced (Xu, 2002). Xu (2009) further identifies two important sources of articulatory constraints for F0 production in speech: the maximum speed of pitch change, and the coordination of laryngeal and supralaryngeal movements (see Section 2.3). It is known that articulatory transition from one pitch level to another takes a discernable amount of time due to the limits on speed (Xu & Sun, 2002). Therefore, it would be more feasible to discuss time instead of speed. As Zhang (2002, 2007) points out, any sequence of insufficient duration would be ruled out to be a contour tone licensor. If a tone bearing unit is not long enough, the contour tone has to be reduced to satisfy this insufficient duration. As the only complex contour tone, T3 is the longest tone in Mandarin and has long been analyzed as preserving its concave contour ([214], three tonal targets) in word-final position. However, the empirical evidence in connected speech indicates that is not the case.

The study of Xu (1997) illustrates that the word-initial T3 only has a falling contour before the middle point except when it is in /T3.T3/, where T3S occurs and the first tone is actually T2. The shape of /T3.T3/ is expectedly similar to the shape of /T2.T3/ (see Fig 6c, p. 75). The study also shows that none of the T3 contours shows the rising shape even when T3 is in word-final position. Instead, only the falling contour of T3 is preserved (see Fig 3 and 6). Recall traditional analysis views a concave T3 as only occurring word-finally and has argued this indicates Beijing Mandarin is right-dominant (see Chao, 1968; Yip, 1980; Lin, 1983; Hashimoto, 1987). The data of Xu (1997) from connected speech suggests differently: T3 is always reduced to a half T3 regardless of its position. In consequence, we argue that Half T3S not only occurs word-initially but word-finally as well.

This phenomenon of a contour tone being reduced exists not only in Beijing Mandarin. If a language does not have a long enough unit to bear a contour tone, different strategies are adopted to resolve the conflict between a sharp pitch excursion and a short duration. Some languages flatten the contour completely (Xhosa), some partially (Pingyao, a Chines dialect), some lengthen the rime (Mitla Zapotec) (Zhang, 2002). The existence of Half T3S in Beijing Mandarin suggests that this language adopts flattening the contour partially to satisfy the insufficient tone-bearing ability. Therefore, I suggest that there is no concave tone in Beijing Mandarin in connected speech; rather, there are only simple contour tones (i.e., HL, LH, two tonal targets). I formalize this restriction on tone-bearing unit as Tonal-DUR (5), and limit the number of tonal targets allowed on disyllables to 4 in Mandarin. That is, a bi-tonal sequence is allowed only when there are no more than 4 tonal targets in Mandarin. This markedness constraint could vary from language to language.

(5) Tonal-DUR: "the maximal tonal targets that a disyllabic sequence can achieve"
A tone's input tonal targets=$t^1t^2t^3...t^n$, a disyllabic sequence=t^{n*2}
A sequence is allowed iff $n*2\leq 4$ in Mandarin.

This time-driven concept is nothing novel. There is ample evidence supporting the physiological claim that only certain articulatory targets can be achieved within allowable time. For example, the studies in Herman et al. (1999), Halle (1994) among others, suggest that there is a finite time between the muscle activities in the larynx and the corresponding F0 changes. Moreover, Lin (1992) proposes a durationally-driven theory and argues that the timing of a toneme is about 1/3 of a tone. For example, T1 is represented as HHH, and T3 is represented as LLM. Lin limits the number of tonemes per tone to three due to the fact that the most complex tones that have been found cross linguistically are the concave tones, which are the combinations of three tonemes, i.e., HLH, HMH. She proposes Tone Reduction Principle (TRP) in connected speech: "reduce a tone by one toneme iff it is immediately followed by another tone within the same prosodic foot" (p. 114). In other words, a bi-tonal sequence allows 4 tonemes in connected speech after TRP application. This TRP triggers both Half T3S and T3S and T3S occurs because of dissimilation.

As a working hypothesis, this study, following Lin (1992), assumes that only four tonal targets are allowed per bi-tonal sequence. Assuming Tonal-DUR is 4 in MC has an essential implication: it triggers Half T3S and T3S. The data from Xu (1997) suggests that T3 is always reduced to a half T3 ([21]) even word-finally while the tonal shapes of other tones remain intact. This shows that a tone-bearing unit in Mandarin fails to bear a sequence of one concave tone (e.g., MLH) but only allows simple contour tones (e.g., HL, LH). That is why Half T3S occurs. Further, if Mandarin cannot afford a sequence of one concave tone, it is not unexpected that a sequence undergoes sandhi if it has adjacent concave tones (i.e., T3S).

It seems that Tonal-DUR could account for question (a) and (b) in the Introduction. T1 is a level tone, which means one tonal target in this study; both T2 and T4 are simple contour tones, which both have two tonal targets per tone. Therefore, a sequence of /T2.T2/ or /T4.T4/ satisfies Tonal-DUR by having four tonal targets at most. Based on the assumption that T3 is always reduced to a half T3 in connected speech no matter whether it appears word-initially or word-finally, Tonal-DUR also provides an explanation for Duanmu's (2007) question why sequences such as /T4.T3/ (H<u>L.L,</u> [51.21]) are legal. This is because there are only four tonal targets in such sequences, which satisfies Tonal-DUR.

3.2 OCP-Cons

Another issue of interest here, however, is more than why other sequence of identical tones survive, but rather what is the precise sandhi output, and how so. If a sequence of /214.35/ becomes /21.35/ (i.e., Half T3S) after the application of Tonal-DUR, a sequence of /214.214/ would expectedly surface as /21.21/. However, Tonal-DUR alone cannot account for the reported output, /35.21/. The next question, also the most important question, is why the sandhi tone is T2.

As discussed above, Moisik et al (2010) show that T3 employs two mechanisms to realize its low pitch targets in producing the citation form: constriction facilitated by larynx raising, or larynx lowering without constriction, while F0 in other tones is correlative to larynx vertical movement only. Namely, the constriction gesture is only found in T3 production particularly during low pitch targets. Therefore, T3 is the only tone that has constriction, which separates T3 from the other tones. Under an AP account, physical properties constrain phonological properties. It is thus reasonable to assume that this distinctive articulatory feature of T3 is responsible for T3 variations provided that this is the principal articulatory difference between T3 and the other tones. Moreover, even though OCP does not trigger T3S, the concept of "no identical units" still can play a role if it is linked to the crucial articulatory distinction between T3 and the other tones. If we define OCP in terms of articulatory gestures, we can propose a markedness constraint on identical gestures focusing on the constriction component of the tones while ignoring larynx height, illustrated in (6).

(6) OCP-Cons: "two consecutive laryngeal constrictions are prohibited"
Let T^1, T^2 be two tones and g^1, g^2 be two laryngeal constriction gestures of T^1, T^2 respectively. The sequence of / $T^1.T^2$ / is prohibited iff $g^1 = g^2$

The concept of defining OCP in an articulatory view is in accordance with Gafos (2002), which considers OCP as "overlapping segments with identical oral gestures are prohibited" (p. 295). In this study, we re-examine it from the perspective of tone production: laryngeal gestures. Like other constraints, OCP-Cons can be language specific. For instance, English can have two adjacent laryngeal constrictions but Mandarin cannot. An independent fact is that consonants are voiceless in Mandarin but voiced in English; however, English allows consonant clusters but MC does not. The maximal syllable structure in Mandarin, as Lin (2001) summarizes, is CGVX (C=consonant, G=glide, V=vowel, X=nasal or vowel) (p. 30). That is, no consonant clusters either in onsets or in codas in Mandarin, i.e., no adjacent constrictions.

Note that OCP-Cons implies two resulting alternatives: either g^1 or g^2 is deleted, leaving only one constriction gesture in a bi-tonal sequence. We have argued that constriction is found in producing the low tonemes. The articulatory gestures of /21.21/ are thus represented as [constriction + larynx raising] + [constriction + larynx raising]. The question is which tone is deleting its constriction feature? Assuming the first constriction gesture of [21] is eliminated, which leaves the first [21] a raising movement. Among the four tones, only T2 is a raising tone. The application in (7) explains the process and illustrates why the sandhi tone has to be a T2.

(7) OCP-Cons Application in T3S

	T3	+	T3
Gesture:	[constriction + larynx raising]		[constriction + larynx raising]
OCP Application:	[larynx raising]	+	[constriction + larynx raising]
Output:	T2+T3		

It seems right so far. However, another possible output arising from (17) is deleting g^2, leaving [constriction + larynx raising] + [larynx raising]. The sandhi tone still is a rising tone, but the output is /T3.T2/ instead of /T2.T3/. So, why the output is not /T3.T2/ since it seems also satisfying OCP-Cons, i.e., the question (c) in Introduction?

We have been arguing so far to incorporate tonal gestures into segment gestural model. If tones behave like consonants suggested in Gao (2008, 2009a, 2009b), one possibility is that the two constriction gestures undergo assimilation: the articulation place of the first tone is assimilated to the second tone in T^1T^2, similar to consonant clusters in English (see Browman & Goldstein, 1990). However, in the sense of AP, a gesture is not deleted, but rather hidden, which still implies two underlying constriction gestures come together. Another possibility is that a sequence of /T3.T3/ can be viewed as a cluster of two identical consonants. In Gafos's (2002) discussion of MCA, the coordination between two consonants with identical oral gestures ([T^{ə}T]) deviates from the default overlapped coordination relationship, CC-COORD. Such deviation would not result in an acoustic release if in an overlap relation. As a result, homorganic clusters employ non-overlap relation and insert a schwa, which violates the default overlapped CC-COORD in MCA and suggests that OCP ranks higher than CC-COORD.

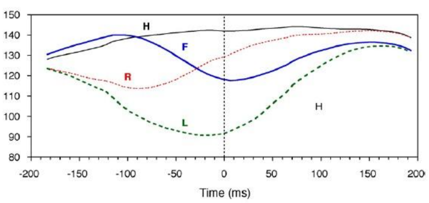
Assuming that there is a default template in Mandarin with respect to tonal coordination relationship, which is violable just like in MCA, would that predict it is g^1 losing constriction gesture and suggest a right output form? In the next section, we will discuss this possibility.

3.3. Mandarin TT-COORD

3.3.1. The bi-tonal relationship

We will start with the optimal citation forms of Mandarin tones. In many cases, the observed F0 contours are "far from" optimal in connected speech (Xu & Wang, 2001, p. 324). Fig (8) below depicts a bi-tonal alternation, drawn from the study of Xu (2009, p. 909). It shows the average F0 contours of a /ma.ma/ sequence with T1 to T4 as T^1 when T1 (H) is T^2 in connected speech.

(8) /ma.ma/ sequence of T1/2/3/4+T1 (H/R/L/F+H)



The Fig. (8) suggests two points. First, the contour of the second tone in a disyllabic sequence is significantly influenced by the offset of the preceding tone: T^2 (H) takes a very long time to reach its high pitch value. In contrast, the anticipatory variations are found to be much smaller. Based on this evidence, Xu (1997) argues that Mandarin has a greater carryover than an anticipatory effect in terms of both magnitude and temporal span (also see Shen, 1990). Second, Fig. (8) shows that the transition between adjacent tones is quite long and that "happens mostly during the H tone itself" (Xu, 2009, p. 908). It takes almost half of T^2's duration before T^2 (H) reaches its "canonical" pitch value as produced in isolation, around 125-130 HZ (see Xu 2009, Fig2a). The long transition is not tone-specific, i.e., H as T^2, but rather quite universal across tones (see Xu, 1997). The long transition thus suggests a shorter duration of T^2 because the transition happens mostly during T^2. In other words, T^2 should be obviously shorter than T^1 when reducing the long transition.

Based on these two points, the coordination relationship between the adjacent tones in MC is referred to as TT-COORD in (9). Transition refers to the dynamic pitch movement between the offset of T^1 and the onset of T^2. To make a stronger point, I argue that this TT-COORD relationship, representing such a coordination relationship in (8), is the default coordination relation in MC.

(9) Inter-tonal coordination: Two tones T^1, T^2 are coordinated with some coordination relation as below.
TT-COORD=ALIGN (T^1, OFFSET, TT, C-CENTER)

(10) Definition: MAX_T: if T^1T^2, preserve all underlying tonemes of T^1

Moreover, consider the bonding strength. C^1 is reduced or assimilated to C^2 in English CC (Browman & Goldstein, 1986). In that case, C^2 has more articulatory strength than C^1. By contrast, the contour of T^2 is significantly influenced by the offset of T^1 and is mostly covered up by T^1 in Mandarin. In consequence, it is T^1 that has more articulatory resistance and thus has more influence on T^2 than vice versa. This coordination relationship requires that T^1 has more articulatory strength and thus is more resistant to change. The definition of TT-COORD is thus shown above in (10). I formalize the articulatory strength within a bi-tonal sequence as a positional faithfulness constraint, MAX_T. It requires that T^1 has more resistance to change and thus all tonal tonemes of T^1 should be intact. From an articulatory point of view, this preservation implies that all articulatory features involved when producing tonemes in T^1 should be intact, which includes two aspects: the larynx vertical movement and the larynx constriction. If this MAX_T constraint, which is projected by TT-COORD, is violable, similar to CC-COORD in MCA, it would suggest that T^1 would lose its articulatory gestures. In this case, it is g^1 will be deleted when demanded. We will show this is the case in an OT fashion in Section 4.

3.3.2. Implication

The proposed bi-tonal coordination relationship in (9) however, contradicts traditional analyses: a concave T3 in Mandarin only occurs on a regularly stressed utterance-final syllable; non-finally it is realized as 21 (Chao, 1930; Zhang, 2002; Yip, 1989). However, Lin (1992, 1994, 2001) and Duanmu (1999) have challenged the theory of final lengthening in Chinese. Lin argues for the initial stress in Mandarin and predicts that disyllabic expressions in Mandarin can be heavy-heavy or heavy-light but not light-light or light-heavy. The empirical evidence measuring the length of the Mandarin disyllabic rime corroborates this argument. The phonetic analysis in Lin, Yan and Sun (1984) shows that the second syllable in a Mandarin disyllabic sequence is either equal to or shorter than the first syllable. Moreover, according to Wang and Wang (1993), it is not the final syllable that is the longest but the initial syllable when an expression is read in a carrier sentence. Duanmu (1991) argues the phonetic facts from Wang and Wang clearly show that the initial syllable is "the most prominent overall" (p.9). In this study, I propose TT-COORD based on phonetic facts in connected speech, and suggest that this supports Lin (1994) and Duanmu's (1991) claims that Mandarin is not right-dominant.

Another issue worth pointing out is that under an AP account, consonant clusters relation (CC-COORD) requires maintaining the recoverability of the two consonants, which is to satisfy a minimal requirement that the produced gestures are generally recoverable by other member of the speech community (Browman & Goldstein, 2000; Goldstein, et al, 2006). In the tone coordination relationship (TT-COORD), we have to take perceptual cause into consideration. As this study views a tone including a string of tonal targets, previous studies suggest that Mandarin tone can be retrieved with reasonable success from rather brief fragments of a syllable's beginning (Lee, 2000; Whalen & Xu, 1992). The main findings in Lee (2000) are that Mandarin listeners are able to correctly identify whether the tone of the target syllable bears a high- or low-onset tone (H, F vs R, L) with 20-40 ms of F0 input from the target syllable. When tones have similar onset pitch, subjects can correctly identify them about 70 ms after the voice onset. Thus, it is possible to leave aside the later portion when the listeners are concerned. In other words, MC has a positional preference for not deleting a tonal target that is tonal-initial.

The faithfulness to the perception, following McCarthy (2008), is thus named Positional $MAX_{initial}$ constraint in this paper, which requires to preserve the left-most toneme. Note that the difference between the two faithfulness constraints proposed in this section is that MAX_T is projected by a bi-tonal coordination relationship in MC to preserve the whole tone, while $MAX_{initial}$ is supposed to preserve the left-most toneme within a tone for a perceptual cause.

(11) Positional $MAX_{initial}$: "a tone's initial tonal target has to be preserved"
Let a tone's input tonal targets=$i_1i_2i_3..i_n$ and output= $o_1o_2o_{3..}o_m$
Assign one violation mark for every i_x if i_x is tonal-initial and there is no correspondent o_Y

To sum up, we explained T3 variations in Beijing Mandarin in an articulatory fashion and proposed several constraints in this section. We proposed that Tonal-DUR triggers Half T3S and T3S for a limited ability of tone-bearing unit in Beijing Mandarin, which rules out /T3.T3/ but keeps intact sequences of other identical tones (the question (a)). We also summarized the articulatory constraints projected by tone production and proposed to define OCP in an articulatory fashion to analyze T3S: OCP-Cons, which disallows adjacent laryngeal constriction in MC (the question (b)). Another two positional faithfulness constraints, one projected by TT-COORD one by perceptual cause, MAX_T and $MAX_{initial}$, respectively, are proposed as well. In the following section, we will discuss further how the interaction among these articulatory constraints determines the output of MC tonal coarticulation (the question (c)).

4. Ranking Constraints

The primary goal of this study is to illustrate Mandarin T3 variations (T3S and Half T3S) with constraints developed under AP. To do so, the projected constraints listed above will also compete against each other in an approach of Optimality Theory (OT, Prince & Smolensky, 1993), accounting for the optimal surface output.

The first deviation from the default TT-COORD is seen in Half T3S, where T3 is reduced to a half T3 to satisfy Tonal-DUR. In the rows of an Optimality Theory tableau (12) below, we compare candidate (12a) with a concave T3 to (12b) where T3 undergoes Half T3S word-initially. As ranked in (12), the candidate (a) violates Tonal-DUR for preserving the concave tone. The candidate (b) violates Positional MAX_T provided that T3 as T^1 in T^1T^2 undergoes change. Since (12b) is the output, we infer that the durational markedness constraint, Tonal-DUR, dominates the faithfulness constraint concerning articulatory strength, MAX_T, that is projected by the default TT-COORD. The fact that MAX_T, further TT-COORD, is violated implies that T^1 would lose its articulatory gestures. In the case of Half T3S, T^1 loses its last toneme, i.e., the raising larynx movement from the perspective of articulation.

(12) Mandarin Half T3S: avoidance of Tonal-DUR; Inferred ranking: Tonal-DUR >> MAX$_T$

Base:T3 (214) .T2 (35)	*Tonal-DUR	MAX$_T$
a. 214.35	*!	
b. ☞21.35		*

Further, we compare the two faithfulness constraints in Tableau (13) to examine another possible output, the candidate (a). In this form, it violates both MAX$_T$ and MAX$_{initial}$ since T^1 changes to [14] and loses its left-most toneme. In contrast, the candidate (b) only violates MAX$_T$ and surfaces as the actual output. We thus infer that MAX$_{initial}$ >> MAX$_{T.}$

(13) Mandarin Half T3S: avoidance of MAX$_{initial}$; Inferred ranking: MAX$_{initial}$ >> MAX$_T$

Base:T3 (214) .T2 (35)	*MAX$_{initial}$	MAX$_T$
a. 14.35	*!	*
b. ☞21.35		*

Although the ranking shown above does not infer the ranking directly between Tonal-DUR and MAX$_{initial}$ since the output (/21.35/) of Half T3S satisfies both constraints, we have argued that it is Tonal-DUR triggers Half T3S. It is reasonable to assume that this markedness constraint should rank above any other constraints. Therefore, we speculate that the ranking should be Tonal-DUR >> MAX$_{initial}$ >> MAX$_T$ (with more evidence coming up concerning T3S).

As we have argued that TT-COORD only accounts for the default template (e.g., Fig. (8)) where no sandhi involves. That is why it cannot account for deviations such as Half T3S, where T^1 loses its partial articulatory features. Considering the default template is violated by Half T3S, it is not unexpected that TT-COORD would be violated in another deviation that is even worse, which is seen as T3S: T3 changes drastically to another tone (i.e., T2) instead of losing partial values. Tableau (14) below compares two candidates concerning this alternation. The candidate (a) in (14) violates Tonal-DUR by having five tonal targets even T^1 goes to a half T3 word-finally. The candidate (14b) incurs the violation of MAX$_T$ because T^1 changes. If the ranking between Tonal-DUR and MAX$_T$ is not quite straightforward in Tableau (12) and (13) concerning Half T3S, it is clear now that Tonal-DUR indeed has a higher rank, suggesting the faithfulness constraint MAX$_T$ also ranks low and is violable in T3S.

(14) Mandarin T3S: avoidance of Tonal-DUR; Inferred ranking: Tonal-DUR >> MAX$_T$

Base: T3 (214) .T3 (214)	*Tonal-DUR	MAX$_T$

a. 214.21	*!	
b. ☞35.21		*

T3S, as we have argued, is also trigged by Tonal-DUR, leaving four tonal targets in /21.21/. That is when another crucial constraint comes into effect, OCP-Cons, which to rule out /21.21/. Tableau (15) enumerates an alternative that is to be compared to the actual output in (15b) with OCP-Cons. The candidate (a) incurs a violation of OCP-Cons for having adjacent low tonal domains, where have adjacent laryngeal constrictions. The candidate (b), however, violates the positional $MAX_{initial}$ twice since the two left-most tonemes of [214] changes to [35]. Despite the multiple violations of $MAX_{initial}$, the candidate (b) still surfaces as the output, suggesting OCP-Cons comes in a higher rank than $MAX_{initial}$, categorically.

(15) Mandarin T3S: avoidance of OCP-Cons; Inferred ranking: OCP-Cons >> $MAX_{initial}$

Base: T3 (214) .T3 (214)	*OCP-Cons	$MAX_{initial}$
a. 21.21	*!	
b. ☞35.21		**

So far, we have shown that Tonal-DUR dominates the two faithfulness constraints in Half T3S. Also, the two markedness constraints dominate two faithfulness constraints in T3S respectively, Tonal-DUR >> MAX_T and OCP-Cons >> $MAX_{initial.}$ From the Tableau (16) below, it can be seen that OCP-Cons dominates the faithfulness constraint, MAX_T, that is projected by TT-COORD. As we have shown that MAX_T is violable in Half T3S, resulting in that T^1 loses its raising feature. In the case of T3S, where both Tonal-DUR and OCP-Cons >> MAX_T and OCP-Cons demands that either T^1 or T^2 deletes a constriction gesture, it is reasonable to infer that should be T^1 because otherwise MAX_T is non-violable. Now it is clear that the application in (17) is correct, and the question (c) is answered.

(16) Beijing Mandarin T3S avoidance of OCP-Cons; Inferred ranking: OCP-Cons >> MAX_T

Base: T3 (214) .T3 (214)	*OCP-Cons	MAX_T
a. 21.21	*!	*
b. ☞35.21		*

Further, the Tableau (27) below combines the four constraints we have discussed so far. It is rather straightforward that the candidate (a) in (27) violates Tonal-DUR and the candidate (b) incurs a violation of the OCP-Cons. However, if $MAX_{Initial}$ dominates MAX_T in T3S as argued in Half T3S, the candidate (d) should surface as the output since it only violates $MAX_{Initial}$ once while the other two candidates violate it twice. Nevertheless, there is no tone of [14] value in Mandarin, and as a result, a sequence of /14.35/ is illegal in this language. It is not in accordance with the economic principle if a language needs to generate a new tone out of a sandhi process that is not underlying specified. Therefore, the candidate (d) can be eliminated.

(17) Beijing Mandarin T3S avoidance of Tonal-DUR and OCP-Cons; Inferred ranking: Tonal-DUR >> OCP-Con >> $MAX_{Initial}$, MAX_T

Base: T3 (214) .T3 (214)	Tonal-DUR	OCP-Cons	$MAX_{Initial}$	MAX_T
a. 214.214	*!			
b. 21.21		*!		*
c. 21.35			**	*
d. 14.35			*	*
e. ☞35.21			**	*

With respect to the candidate (c) and (e), it seems that they are equally worse. However, remember that MAX_T is violable suggests that T^1 is in a disadvantage position in terms of preserving articulatory features in T^1T^2. Also, the application of OCP-Cons not only prohibits adjacent constriction gestures, it also predicts the sandhi tone has to be a rising tone in Section 3.2. Nevertheless, the candidate (c)'s sandhi tone is a falling tone. The two points lead to why the candidate (e) wins out against the candidate (c) even though they both violate $MAX_{initial}$ and MAX_T.

Despite we cannot infer the ranking between Tonal-DUR and OCP-Cons from Tableau (17), considering all sandhi variations are triggered by the durational-driven constraint, Tonal-DUR should have the highest ranking and dominate all other constraints. Therefore, the ranking in T3S is drawn as Tonal-DUR >> OCP-Cons >> $MAX_{initial}$, $MAX_{T.}$ The ranking between $MAX_{initial}$ and MAX_T is unclear in T3S, but we have shown that $MAX_{initial}$ >> MAX_T in Half T3S. Therefore, we propose that in Beijing Mandarin the four constraints should have a ranking as Tonal-DUR >> OCP-Cons >> $MAX_{initial}$ >> MAX_T.

To sum up, this section further discussed several articulatory and perceptual constraints and ranked them in OT. The results show that the two markedness constraints rank above the two faithfulness constraints, and that Tonal-DUR has the highest ranking in Beijing Mandarin. The ranking between $MAX_{initial}$ and MAX_T is somewhat unclear in T3S for now given no further evidence, will be considered for future study.

5. Conclusions

The most important merit of this study is that it shows central hypotheses under AP model are also applicable to tone sandhi analysis and is phonologically economic. The present analysis could be further tested however, e.g., in Shanghai dialect. Differing from (Beijing) Mandarin in that most regular syllables keep their underlying tones, Shanghai dialect has the initial syllable determining the whole tonal pattern and has much shorter syllable duration (see Duanmu, 1994). By doing so, we would have a better understanding of these AP constraints could account for more tonal alternations across Chinese dialects.

References

Browman, C, & Goldstein, L. (1992). Articulatory Phonology: An overview. *Phonetica,* 49, 155-180.

Duanmu, S. (1994). Against Contour Tone Units. *Linguistic Inquiry*.25 (4). 555-608

Duanmu, S. (2004). Left-headed feet and phrasal stress in Chinese. *Cahiers de linguis tique - Asie orientale*, vol. 33. 65-103.

Duanmu, S. (2007). *The Phonology of Chinese*. Oxford: Oxford University Press.

Gafos, A. (2002). A Grammar of Gestural Coordination. *Natural language and Linguistic theory*, 20, 269-337.

Gao, Man. (2008). *Mandarin tones: an articulatory phonology account*. PhD Dissertation, Yale University.

Herman, R., Beckman, M., & Honda, K. (1999). Linguistic Models of F0 use, physiolo gical models of F0 control, and the issue of "mean response time". *Language and Speech*, 42: 373-399.

Honda, K., Hirai, H, Masaki, S., & Shimada, Y. (1999). Role of vertical larynx move ment and cervical lordosis in F0 control, *Language Speech*, 401-411.

Lee, C-Y. (2000). Lexical tone in spoken word recognition: a view from Mandarin Chi nese. PhD dissertation. Brown University.

Lin, Hua. (1992). On the nature of Mandarin tone and tone sandhi. PhD disserta tion. University of Victoria.

Lin, Hua. (2001). Stress and the distribution of the neutral tone in Mandarin. *Chinese phonology in generative grammar*, 139-162.

McCarthy, J. J. (2008). *Doing optimality theory: Applying theory to data*. Malden, MA, and Oxford, UK: Blackwell.

Moisik, S. R., Lin, H., & Esling, J. H. (2010). An investigation of laryngeal behavior during Mandarin tone production using simultaneous laryngoscopy and laryngeal ultrasound. *In Proceedings of the 9th Phonetics Conference of China.*

Wang, Gaowu., & Kong, Jiangping. (2010). The relation between laryngeal height and F0 during the four tones of Mandarin in X-ray movie. Proc. ISCSLP, Taiwan, 335 338.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61–83.

Xu, Y. (2009). Timing and coordination in tone and intonation-An articulatory-functional perspective. *Lingua*, 119, 906-927.

Yip, Moria. (2002). *Tone*. Cambridge: Cambridge University Press

Zhang, J. (2002). *The Effects of Duration and Sonority on Contour Tone Distribution — a Typological Survey and Formal Analysis*. New York: Routledge.