Binarity and Focus in Prosodic Phrasing: New Evidence from Taiwan Mandarin

Shu-hao Shih
Rutgers University

1 Introduction

This paper makes novel claims and presents new evidence for the binarity of prosodic phrases, alignment of focused elements, and their interaction. Several authors have argued that there are binarity restrictions at the level of the prosodic phrase (e.g. Prince 1980, Ito & Mester 1992, Selkirk 2000). In this paper, I provide new support for this claim from Taiwan Mandarin (TM). I argue that prosodic phrases must be decomposed into Minor Phrases (MIPs) and Major Phrases (MAPs) (Selkirk *et al*. 2004), and that both levels can have minimal and maximal binarity restrictions. Crucially, MAPs must be binary in TM. Moreover, I argue – after Féry (2013) – that focused elements can require *both* left and right edge alignment of a constituent (in TM, a MIP). Finally, I show that requirements on binarity and focus can interact in striking ways – a binarity restriction on MAPs prevents alignment to the right edge of focused elements in specific environments.

This paper is organized in the following manner. Section 2 presents the methodology used to investigate default and focal phrasing in TM. Sections 3 and 4 report the results and offer a formal analysis in Optimality Theory (Prince & Smolensky 1993) for the default and focal phrasing patterns, respectively. Section 5 concludes this paper.

2 Methodology

2.1 *Experiment design* The goal of this experiment was to investigate both default and focal prosodic phrasing in TM. In the TM numeral system, /wuL/ '5' and /tɕjouL/ '9' bear a lexical low tone. Other numerals bear a non-low tone, so there is no tone sandhi when they are adjacent. Numbers greater than ten are disyllabic words, so only /wuL/ '5' and /tɕjouL/ '9' are eligible for testing the phrasing pattern.

For the default phrasing, the five-word string /wuL wuL wuL wuL wuL/ '55555' and the six-word string /wuL wuL wuL wuL wuL wuL/ '555555' were used to examine the prosodic phrasing patterns. As the numeral strings are syntactically flat, prosodic phrasing is not influenced by syntactic structure (Chen 2000: 368). The reason that /wuL/ '5' was chosen to test the default phrasing pattern was that the onset of /tɕjouL/ '9' is an affricate, so the consonant closure might be confused with a pause. As discussed in Section 3, pause is an important diagnostic for phrase boundaries.

For focal phrasing, /tɕjouL/ '9' was used to test the phrasing pattern. Each word in the default string was replaced by /tɕjouL/ '9' one at a time, as illustrated in (1). The replacement of each position ensured that all the possible focal phrasing patterns were elicited.

* For helpful comments on various versions of this work, I thank Chelsea Sanker, Paul de Lacy, Yi-Hsun Chen, Ryan Bennett, Nick Danis, Luca Iacoponi, Gwendolyn Rehrig, Vartan Haghverdi, Hazel Mitchley, Eileen Blum, and Christopher Oakden. I also thank audiences at the 2016 Annual Meeting on Phonology (AMP 2016), the 34th West Coast Conference on Formal Linguistics (WCCFL 34), the 2016 Phonology in the Northeast (PhoNE 2016), and the Rutgers Phonology and Phonetics Research Group (PhonX). All remaining errors are mine.

Proceedings of *AMP 2016*

(1) Stimuli

Default strings	Focal strings
/wuL wuL wuL wuL wuL/ ‘55555’	/**tɕjouL** wuL wuL wuL wuL/ ‘**9**5555’
	/wuL **tɕjouL** wuL wuL wuL/ ‘5**9**555’
	/wuL wuL **tɕjouL** wuL wuL/ ‘55**9**55’
	/wuL wuL wuL **tɕjouL** wuL/ ‘555**9**5’
	/wuL wuL wuL wuL **tɕjouL**/ ‘5555**9**’
/wuL wuL wuL wuL wuL wuL/ ‘555555’	/**tɕjouL** wuL wuL wuL wuL wuL/ ‘**9**55555’
	/wuL **tɕjouL** wuL wuL wuL wuL/ ‘5**9**5555’
	/wuL wuL **tɕjouL** wuL wuL wuL/ ‘55**9**555’
	/wuL wuL wuL **tɕjouL** wuL wuL/ ‘555**9**55’
	/wuL wuL wuL wuL **tɕjouL** wuL/ ‘5555**9**5’
	/wuL wuL wuL wuL wuL **tɕjouL**/ ‘55555**9**’

Two frame sentences were used in this experiment, as in (2). The purpose was to test the default phrasing pattern and see whether the default phrasing pattern is different under the condition of contrastive focus. In the three environments shown below, the default strings (i.e. 55555, 555555) were placed in the first and second environments, and the focal strings (e.g. 95555, 955555) were put in the third environment. The first sentence is a statement. The second sentence negates the previous statement by first repeating the default string in the first sentence and then asserting the correct focal string.

(2) Frame sentences

a. [tjɛnHL xwaHL xauHL maL ʂɻ̩HL ___①___.] ‘The telephone number is ___.’

b. [puHL, puMH ʂɻ̩HL ___②___, ʂɻ̩HL ___③___.] ‘No, it’s not ___. It’s ___.’

One telephone and two dialogue icons were used in this experiment, as illustrated in Figure 1. The purpose was to not only facilitate elicitation but also help participants become familiar with the context. The dialogue on the left contains the default string while the one on the right the focal string. All the graphs were presented on a PowerPoint slide. An animation was added to the slide so that the dialogue on the right appeared later than the dialogue on the left.

Figure 1: Experiment setup.

In short, the stimuli, frame sentences, and graphic icons ensured that a contrastive focus was elicited.

There were six recording sessions in this experiment. The first three sessions were for the five-word strings, whereas the last three sessions were for the six-word strings. Participants read 45 stimuli in the first three sessions, and 54 stimuli in the last three sessions, yielding a total of 99 tokens per speaker. Moreover, 44 fillers with other numbers (20 for the five-word string and 24 for the six-word string) were also included to divert speakers’ attention from the tone sandhi patterns. The order of the stimuli was pseudo-randomized in each session. The order of the recording sessions was also counter-balanced.

2.2 *Participants* Two female and three male native TM speakers participated in the experiment. Their ages ranged from 21 to 32 years old. All had recently moved to the United States and still communicated in TM on a daily basis. None of the participants had linguistic training or a history of speech impairments. They were naïve as to the goal of the experiment.

2.3 *Procedure* The experiment was performed at the Phonology and Field Research Laboratory (Phonolab) at Rutgers University. Participants were recorded while sitting in a sound-attenuated booth and wearing an AKG C420 head-worn microphone with a behind-the-neck headband in order to keep the microphone at a constant distance from the mouth. The microphone was connected to an ART MPA Gold pre-amplifier, with output to an M-Audio Delta 1010LT sound card. The recording was done using GoldWave v6. 10 at a 44.1k Hz sampling rate and 16-bit quantizing rate in mono.

Participants were presented with several slides with a telephone icon and two stimuli on a computer screen. The two frame sentences in traditional Mandarin characters were printed out on a paper. When the dialogue on the left appeared on the slide, participants were asked to read the first frame sentence. When the dialogue on the right came out, participants had to read the second frame sentence. Participants performed the task when they were ready, at a normal conversational speed. Breaks were given after each recording session. Some effort was expended in ensuring that the participants employed their vernacular. Specifically, the investigator engaged the participants in vernacular speech by having conversations with them during the breaks about mundane daily activities. Participants received a training session to familiarize themselves with the task. The recording sessions were conducted individually.

2.4 *Analytical method* Glottal stops/glottalization and pauses were labeled in Praat TextGrids for analyzing prosodic boundaries. The motivation for this is that previous studies have shown that the occurrence of glottal stop/glottalization and pause indicates higher prosodic level boundaries. Cross-linguistically, non-modal phonation (in particular creaky voice) is frequently used to mark prosodic boundaries, either initially and/or finally (Gordon & Ladefoged 2001). This kind of allophonic glottalization has been documented in a range of languages, including English (Pierrehumbert & Talkin 1992, Dilley *et al*. 1996), German (Kohler 1994) and French (Fougeron 2001). For example, English speakers are more likely to glottalize word-initial vowels when those vowels occur at the beginning of a new Intonational Phrase (Pierrehumbert & Talkin 1992, Dilley *et al*. 1996). Pauses are also indicators of Intonational Phrase boundaries (Nespor & Vogel 1987, Selkirk 1984).

Tone sandhi was also used as a diagnostic of prosodic boundaries; it occurs iteratively from left to right when two underlying low tones are adjacent: /L L/ → [MH L]. Two native speakers of TM performed an auditory analysis of the tone sandhi patterns: they coded tokens as either L or MH; their judgments were included in the analysis only when both listeners agreed on the classification; there was 99% inter-rater agreement. Data with disfluency or hesitation was excluded from the analysis.

3 Default phrasing

Three properties provide clear diagnostics for phrase boundaries: (a) a pause indicates a Major Phrase (MAP) boundary; (b) a glottal stop/glottalization marks a Minor Phrase (MIP) boundary; and (c) failure of tone sandhi to apply indicates a MAP boundary, as tone sandhi only occurs within MAPs.

For the five-word string, all participants produced two MAPs of two and three Prosodic Words, primarily {[MH][L]}{[MH.MH][L]}, as in Figure 2 (Note: { } mark MAP boundaries; [ ] mark MIP boundaries). Three of the participants also produced the reverse pattern, {[MH.MH][L]}{[MH][L]}, as in Figure 3. The production frequencies and percentages of the tone sandhi patterns are given in Table 1. Tone sandhi applies within each MAP; the last syllable in each MAP retains a L tone. Except for participant M1, the prosodic structure {[MH][L]}{[MH.MH][L]} was more frequent than the prosodic structure {[MH.MH][L]}{[MH][L]}.

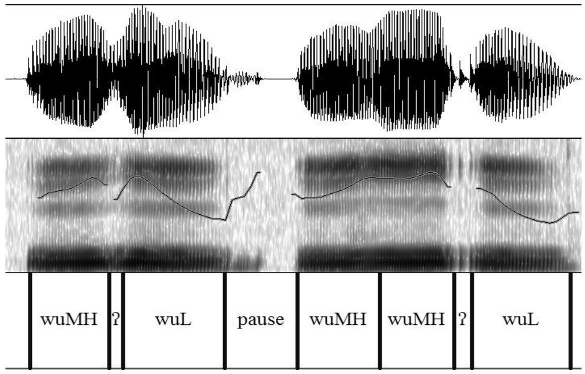


Figure 2: Prosodic phrasing for the five-word string: {[MH][L]}{[MH.MH][L]}.

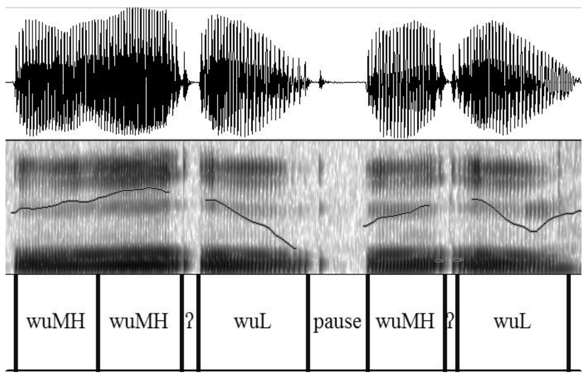


Figure 3: Prosodic phrasing for the five-word string: {[MH.MH][L]}{[MH][L]}.

Participant	{[MH][L]}{[MH.MH][L]}	{[MH.MH][L]}{[MH][L]}
M1	12 (44%)	15 (56%)
M2	20 (77%)	06 (23%)
F1	16 (59%)	11 (41%)
M3	30 (100%)	00 (0%)
F2	28 (100%)	00 (0%)

Table 1: Production frequencies and percentages for the five-word string.

For the six-word string, all participants consistently produced two MAPs of three Prosodic Words (ω) each, with the first two in one MIP and the last one in another MIP: {[MH.MH][L]}{[MH.MH][L]}, as shown in Figure 4. There was no variation in the six-word string. The production frequencies and percentages of the tone sandhi pattern are given in Table 2.

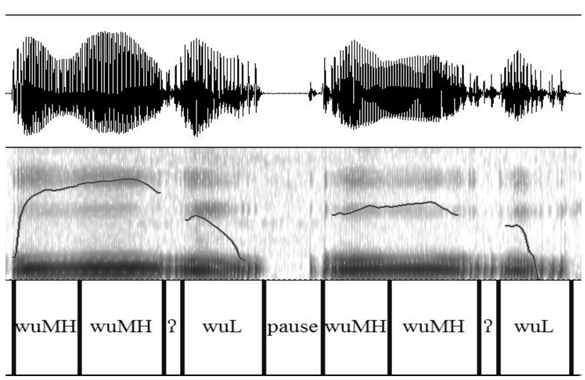


Figure 4: Prosodic phrasing for the six-word string: {[MH.MH][L]}{[MH.MH][L]}.

Participant	{[MH.MH][L]}{[MH.MH][L]}
M1	32 (100%)
M2	36 (100%)
F1	36 (100%)
M3	28 (100%)
F2	30 (100%)

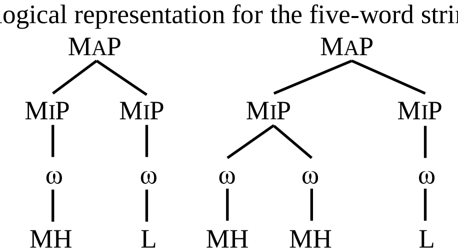
Table 2: Production frequencies and percentages for the six-word string.

Based on the phonetic and phonological diagnostics, I propose that prosodic phrases in TM must be decomposed into MAPs and MIPs. MAPs dominate MIPs. The MIP is the prosodic constituent that immediately dominates the Prosodic Word (ω) in the prosodic hierarchy (Selkirk *et al*. 2004). Since each syllable in the numeral string is a content word, each syllable forms a Prosodic Word, which is directly dominated by a MIP.[1] The whole numeral string is an Intonational Phrase, directly dominating MAPs. Therefore, {[MH]} refers to a MAP that dominates a MIP that dominates a Prosodic Word that dominates a foot that dominates a syllable node that bears the contour tone MH.

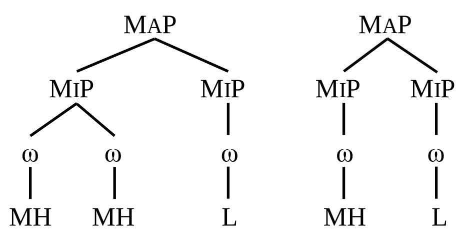
Because of the variation among speakers for the five-word string, I propose there are two prosodic structures that occur in surface forms for the five-word string, whereas there is one that occurs for the six-word string. The phonological representations are provided in (3) and (4) (Note: ω = Prosodic Word, MIP = Minor Phrase, MAP = Major Phrase). Essentially, the phonetic and phonological evidence clearly show that MAPs must be binary, containing two MIPs. MIPs prefer to be binary but can also contain a single Prosodic Word.

(3) Phonological representation for the five-word string

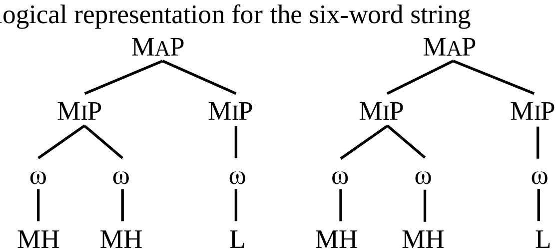
a.



b.



(4) Phonological representation for the six-word string



Below I offer a formal account for the default prosodic phrasing using Optimality Theory. I propose that the asymmetry between MAPs and MIPs is due to constraints on minimal and maximal size of the relevant

[1] According to L. Cheng (1987), two content words cannot be grouped into one Prosodic Word in Mandarin.

prosodic domains and the ranking of these constraints. In Optimality Theory, binarity constraints usually come in separate versions for minimal and maximal binarity (e.g. Mester 1994, Selkirk 2000, Elias-Ulloa 2006, Ito & Mester 2007). Constraints on the minimum and maximum size of prosodic constituents are argued to be part of the universal repertoire.

(5) Constraints on size restriction
a. MAPMAX: Incur a violation for each Major Phrase that dominates more than two nodes.
b. MAPMIN: Incur a violation for each Major Phrase that dominates fewer than two nodes.
c. MIPMAX: Incur a violation for each Minor Phrase that dominates more than two nodes.
d. MIPMIN: Incur a violation for each Minor Phrase that dominates fewer than two nodes.

Additionally, PARSE constraints require strict layering: PARSE-ω requires every Prosodic Word to be dominated by a MIP; PARSE-MIP requires every MIP to be dominated by a MAP. In TM, unparsed Prosodic Words and MIPs never occur in winners. So, both PARSE-ω and PARSE-MIP outrank any antagonistic constraints. Therefore, any candidates with unparsed elements such as *{[ωω]ω} and *{[ωω]}[ω] are eliminated by one of the PARSE constraints.

MAPMAX and MAPMIN must outrank MIPMIN to ensure that MAPs are obligatorily binary. By ranking MAPMAX over MIPMIN, a MAP which contains three MIPs is ruled out, as shown in (6c). Ranking MAPMIN over MIPMIN excludes a MAP with only one MIP, as in (6d). In short, having MAPMAX and MAPMIN dominate MIPMIN guarantees that MAPs are strictly binary.

On the other hand, MIPs can be either unary or binary. This is captured by ranking MIPMAX over MIPMIN. By doing so, a MIP which contains three Prosodic Words fatally violates MIPMAX, as in (6e). Consequently, a MIP with either one or two Prosodic Words is allowed.

(6) Maximal and minimal binarity restrictions on MAPs and MIPs

/55555/	MAPMAX	MAPMIN	MIPMAX	MIPMIN	ALIGN-R (MIP, MAP)
☞a. {[ω][ω]}{[ωω][ω]}				***	* *
☞b. {[ωω][ω]}{[ω][ω]}				***	* *
c. {[ωω][ωω][ω]}	*!			*	*** *
d. {[ωω][ω]}{[ωω]}		*!		*	*
e. {[ωω][ωωω]}			*!		***
f. {[ω][ω]}{[ω][ωω]}				***	* **!

Notice that candidate (6f) has the prosodic structure *{[ω][ωω]}, which is the reverse pattern of {[ωω][ω]} in candidates (6a) and (6b). The prosodic structures *{[ω][ωω]} and {[ωω][ω]} incur equal violations of MIPMIN since both have a MIP with one Prosodic Word. I argue that the ungrammatical candidate *{[ω][ωω]} fatally violates the constraint ALIGN-R(MIP, MAP) twice because the right edge of the first MIP is two Prosodic Words away from the right edge of the MAP, whereas {[ωω][ω]} violates ALIGN-R(MIP, MAP) once.

One remaining question is why there is variation among subjects. For participants M1, M2, and F1, both candidates (6a) and (6b) are optimal outputs, whereas for participants M3 and F2, only candidate (6a) is the optimal output. It should be noted that candidates (6a) and (6b) have the same violation profile in (6); they both violate MIPMIN three times. Thus, the proposed constraints are not sufficient to account for the variation observed in the five-word string. I propose that variation occurs due to competing requirements for the MIP edges to be aligned with Intonational Phrase edges.

(7) Edge alignment between MIP and Intonational Phrase
a. ALIGN-L(MIP, IP):
Align the left edge of every Minor Phrase with the left edge of an Intonational Phrase.
b. ALIGN-R(MIP, IP):
Align the right edge of every Minor Phrase with the right edge of an Intonational Phrase.

(8) Partially ordered alignment constraints

/55555/	Align-L (MiP, IP)	Align-R (MiP, IP)
a. {[ω][ω]}{[ωω][ω]}	7	8
b. {[ωω][ω]}{[ω][ω]}	9	6

Assuming the whole numeral string is an Intonational Phrase, candidates (8a) and (8b) exhibit a conflict in the alignment constraints. (8b) incurs more violations of Align-L(MiP, IP) than (8a) does, because the MiPs are more distant from the left edge of the Intonational Phrase. However, (8a) incurs more violations of Align-R(MiP, IP) than (8b) does, because the MiPs are more distant from the right edge of the Intonational Phrase. For participants with variation, Align-L(MiP, IP) and Align-R(MiP, IP) are partially ordered (see Anttila 2007). Hence, both candidates can be optimal outputs. For participants with no variation, Align-L(MiP, IP) dominates Align-R(MiP, IP). As a result, candidate (8a) is the only winner. No participant has the grammar Align-R(MiP, IP) » Align-L(MiP, IP) because no participant solely produced the prosodic structure {[ωω][ω]}{[ω][ω]}. It is also possible that the variation could be explained by using constraints that align MaP edges with Intonational Phrase edges. Unfortunately, the current data does not provide enough evidence to distinguish the two approaches and developing this line of research would go beyond the scope of this paper.

The proposed constraints and their ranking successfully predict the prosodic phrasing of the six-word string. A tableau for the six-word string is provided in (9). The important point here is that even though the length of the numeral string is changed from five to six, MaPs still maintain their binarity. Candidate (9b) violates MaPMax because the MaP contains three MiPs. Candidate (9c) incurs three violations of MaPMin because each MaP includes only one MiP. Candidate (9d) has two MiPs, but each MiP has three Prosodic Words, thus violating MiPMax. Candidate (9e) is ruled out because the right edge of the first MiP in each MaP is two Prosodic Words away from the right edge of the MaP. Finally, recall that there is no variation for the six-word string. I argue that this is because the alignment constraints which determine the variation observed for the five-word string are not crucial here. They do not distinguish any losing candidates from the winner. As a result, there is only one possible prosodic phrasing pattern for the six- word string.

(9) Constraint ranking for the six-word string

/555555/	MaPMax	MaPMin	MiPMax	MiPMin	Align-R (MiP, MaP)	Align-L (MiP, IP)	Align-R (MiP, IP)
☞a. {[ωω][ω]}{[ωω][ω]}				**	* *	10	8
b. {[ωω][ωω][ωω]}	*!				**** **	6	6
c. {[ωω]}{[ωω]}{[ωω]}		***!				6	6
d. {[ωωω][ωωω]}			**!		***	3	3
e. {[ω][ωω]}{[ω][ωω]}				**	** **!	8	10

4 Focal phrasing

The results for the five-word string are summarized in (10). The focused element in each string forms a MiP itself. For (10a), (10b), (10c), and (10e), the focused element retains its underlying L tone. Notice that when the focused element is the last word, there is free variation. Nevertheless, the focused element forms a MiP on its own. This variation is only found in the data from participants M1, M2, and F1.

(10) Five-word string with contrastive focus
a. {[**ω**][ω.ω]}{[ω][ω]} ‘**9**5555’
b. {[ω][**ω**]}{[ω.ω][ω]} ‘5**9**555’
c. {[ω.ω][**ω**]}{[ω][ω]} ‘55**9**55’
d. {[ω.ω][ω]}{[**ω**][ω]} ‘555**9**5’
e. {[ω.ω][ω]}{[ω][**ω**]} ‘5555**9**’
or {[ω][ω]}{[ω.ω][**ω**]}

(11) Six-word string with contrastive focus
a. {[**ω**][ω.ω]}{[ω.ω][ω]} ‘**9**55555’
b. {[ω][**ω**.ω]}{[ω.ω][ω]} ‘5**9**5555’
c. {[ω.ω][**ω**]}{[ω.ω][ω]} ‘55**9**555’
d. {[ω.ω][ω]}{[**ω**][ω.ω]} ‘555**9**55’
e. {[ω.ω][ω]}{[ω][**ω**.ω]} ‘5555**9**5’
f. {[ω.ω][ω]}{[ω.ω][**ω**]} ‘55555**9**’

For the six-word string, there are six focal phrasing patterns, as illustrated in (11). Three observations can be made based on the focal phrasing patterns. Firstly, the focused element forms a MIP at the beginning of a MAP, as in (11a) and (11d). Secondly, the focused element forms a MIP with the following Prosodic Word when the focused element is at the second or fifth position, as in (11b) and (11e). Finally, the focused element forms a MIP itself at the end of a MAP, as in (11c) and (11f).

As for tonal variations, the focused element retains its L tone in (11a), (11c), (11d), and (11f), whereas the focused element undergoes tone sandhi in (11b) and (11e).

When the position of focus is changed, the phrasing pattern may be different from the default phrasing pattern. Note that the default phrasing pattern is {[ω.ω][ω]}{[ω.ω][ω]} ‘555555’. Of the six patterns, four deviate from the default phrasing pattern, as in (11a), (11b), (11d), and (11e). The rest of the two patterns (11c) and (11f) match the default phrasing pattern.

However, the production frequencies for the five- and six- word strings are complicated among participants. For the five-word string, participants M3 and F2 produced the same focal phrasing pattern as the default phrasing pattern. It seems that they are not sensitive to the existence of contrastive focus. Recall that participants M3 and F2 generated only one default phrasing pattern for the five-word string, as shown in Table 1. However, they produced the focal phrasing pattern for the six-word string. As summarized in Table 3, all the participants produced the focal phrasing pattern for the six-word string, but there are production variations for the five-word string. This fact suggests that participants might be sensitive to the length of the material or have different grammars. I leave it for future work.

	Five-word string	Six-word string
Participant	Default phrasing vs. Focal phrasing	Default phrasing vs. Focal phrasing
M1	Different	Different
M2	Different	Different
F1	Different	Different
M3	*Identical*	Different
F2	*Identical*	Different

Table 3: Comparisons between default and focal phrasing.

In short, the generalization is that the focused element forms a MIP on its own, except when doing so would force the MIP to contain two Prosodic Words (see 11b and 11e). Essentially, focus forces the arrangement of MIPs to deviate from the default pattern so that MAPs can still maintain their binarity.

Féry (2013) argues that focus universally tends to be aligned prosodically with the right or left edge of a prosodic domain. To be specific, a focused constituent is preferably aligned with the right or left edge of either a Prosodic Phrase or an Intonation Phrase. Languages have different strategies to fulfill the alignment constraint, such as syntactic movement, cleft constructions, insertion of a prosodic boundary, and enhancement of existing boundaries. Féry then proposes a family of constraints ALIGN-FOCUS in the framework of Optimality Theory, as shown in (12).

(12) ALIGN-FOCUS (Féry 2013: 691)
a. ALIGN-FOCUS R, ι-PHRASE R (ALIGN-FOC-ι-R):
Align a focus with the right boundary of an Intonation Phrase.
b. ALIGN-FOCUS L, ι-PHRASE L (ALIGN-FOC-ι-L):
Align a focus with the left boundary of an Intonation Phrase.
c. ALIGN-FOCUS R, ϕ-PHRASE R (ALIGN-FOC-ϕ-R):
Align a focus with the right boundary of a Prosodic Phrase.
d. ALIGN-FOCUS L, ϕ-PHRASE L (ALIGN-FOC-ϕ-L):
Align a focus with the left boundary of a Prosodic Phrase.

After Féry (2013), I propose that focal phrasing in TM is governed by constraints aligning the edges of the focused element with the edges of a MIP: ALIGN-L(FOCUS, MIP) and ALIGN-R(FOCUS, MIP). Below I will use the six-word string to illustrate constraint ranking.

(13) Constraints on focus alignment
a. ALIGN-L(FOCUS, MIP):
Align the left edge of the focused element with the left edge of a Minor Phrase.
b. ALIGN-R(FOCUS, MIP):
Align the right edge of the focused element with the right edge of a Minor Phrase.

Crucially, ALIGN-L(FOCUS, MIP) must outrank ALIGN-R(FOCUS, MIP) because in some situations perfect MIP-to-focus alignment is not possible (namely when MAPs would otherwise be ternary or unary): i.e. in {[ω][**ω**.ω]}, not *{[ω][**ω**][ω]} (see 14c) or *{[**ω**]} (see 14d). We can see in this situation that the need for maximal and minimal binarity in MAPs outweighs the need to align the right edge of the focused element with a MIP. Candidate (14b) fatally violates ALIGN-L(FOCUS, MIP) because the left edge of the focused element does not align with the left edge of a MIP. Candidates (14c) and (14d) perfectly satisfy the alignment constraints on focus. However, the MAPs are not strictly binary: the MAP is ternary in candidate (14c), and unary in candidate (14d).

(14) Focus on the second Prosodic Word

/595555/	MAPMAX	MAPMIN	ALIGN-L (FOCUS, MIP)	ALIGN-R (FOCUS, MIP)
☞a. {[ω][**ω**.ω]}{[ω.ω][ω]}				*
b. {[ω.**ω**][ω]}{[ω.ω][ω]}			*!	
c. {[ω][**ω**][ω]}{[ω.ω][ω]}	*!			
d. {[ω]}{[**ω**]}{[ω.ω]}{[ω.ω]}		**!		

Furthermore, both alignment constraints must be active in the language to compel deviation from the default phrasing, as in {[**ω**][ω.ω]}{[ω.ω][ω]}. Essentially, focus can require alignment of both left and right boundaries. Candidate (15a) incurs no violation of ALIGN-L(FOCUS, MIP) and ALIGN-R(FOCUS, MIP), whereas candidate (15b) fatally violates ALIGN-R(FOCUS, MIP).

(15) Focus on the first Prosodic Word

/955555/	ALIGN-L (FOCUS, MIP)	ALIGN-R (FOCUS, MIP)	ALIGN-R (MIP, MAP)
☞a. {[**ω**][ω.ω]}{[ω.ω][ω]}			** *
b. {[**ω**.ω][ω]}{[ω.ω][ω]}		*!	**

The same analysis apples to the situation where the focus is at the third or sixth word, as shown in (16). All the candidates in (16) obey the binarity restriction on MAPs, but (16b) violates ALIGN-L(FOCUS, MIP) and (16c) ALIGN-R(FOCUS, MIP).

(16) Focus on the third Prosodic Word

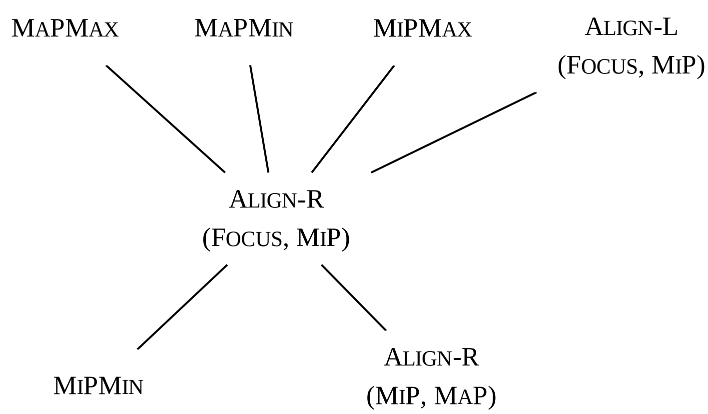
/559555/	ALIGN-L (FOCUS, MIP)	ALIGN-R (FOCUS, MIP)	ALIGN-R (MIP, MAP)
☞a. {[ω.ω][**ω**]}{[ω.ω][ω]}			* *
b. {[ω][ω.**ω**]}{[ω.ω][ω]}	*!		** *
c. {[ω][ω]}{[**ω**.ω][ω.ω]}		*!	* **

In short, this section shows that (a) one focus edge-alignment is prioritized over the other, (b) focus can require alignment of both left and right boundaries, and (c) a MAP still maintains its binarity under the condition of contrastive focus.

Additional constraints are needed to account for numeral strings where the first or second Prosodic Word is focused. With the ranking above, the candidate *{[ω][ω]}{[ω.ω][ω.ω]} will win by virtue of aligning MIP boundaries with the focused constituent (the first or second Prosodic Word), and incurring only three violations of ALIGN-R(MIP, MAP). However, the fatal flaw with this losing candidate is that its second MAP is too large: it contains more than three Prosodic Words – something found in no winning candidate. TM's limit on the maximum size of MAPs is analogous to maximal size restrictions on Prosodic Words, which may contain a maximum of three dependents in some languages (de Lacy 2004, Ketner 2007). A full analysis will not be provided here due to space limitations.

The resulting ranking for TM is summarized in (17). Note that the constraints on participant variation are not included in the diagram. The interaction between binarity and focus is clearly reflected in the ranking hierarchy. When a contrastive focus is involved in prosodic phrasing, the binarity requirement on MAPs is strictly obeyed. The binarity requirement on MIPs is relatively flexible compared to MAPs when focus comes into play. In other words, a MIP could be either unary or binary when it interacts with focus.

(17) Taiwan Mandarin default/focal phrasing ranking summary



5 Conclusion

Many studies have argued that there are binarity requirements at different prosodic levels, such as feet (Prince 1980), Prosodic Words (Ito & Mester 1992), Major Phrases (Selkirk 2000), and so on. However, it is difficult to tell whether higher prosodic levels hold such restrictions because syntax usually interferes and the length of the material matters. The experiment reported here avoids the interference of syntactic structure, allowing phonological restrictions to shine through. Significantly, this paper shows that the binarity restriction holds not only at the MAP level but also at the MIP level: both MAPs and MIPs can require both maximal and minimal binarity.

Féry (2013) identifies languages in which focus aligns with the right edge of a prosodic domain (e.g. Italian, French, Chicheŵa, Northern Biskayan Basque, and Konkani) and with the left edge of a prosodic domain (e.g. Hungarian, Nłeʔkepmxcin, and Georgian). However, the missing case is where focus aligns with both left and right edges of a prosodic domain. Crucially, this paper shows that focus in TM can require alignment of both left and right focus boundaries. This finding thus fills in Féry (2013)'s typological gap.

This paper further demonstrates that one edge-alignment can be prioritized over the other, and that binarity and focus restrictions can interact in striking ways, illuminating the nature of both kinds of conditions on prosodic phrasing. When focus comes into play in prosodic phrasing, the formation of MIPs may be changed but MAPs still maintain their binarity.

Finally, this paper shows that glottal stops and pauses play an important role in identifying prosodic boundaries in TM. In the past, many studies have written about Mandarin tone sandhi and this has been the only evidence cited for prosodic structure (see Shih 1986, 1997, Chen 2000 for discussion). The relevance of glottal stops and pauses are therefore significant in prosodic phonology.

References

Antilla, Arto (2007). Variation and optionality. In Paul de Lacy (ed.) *The Cambridge Handbook of Phonology*. Cambridge University Press. 519–536.

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

Cheng, Lisa (1987). Derived domains and Mandarin third tone sandhi. *CLS* **23**. 16–28.

de Lacy, Paul (2004). Maximal Words and the Māori passive. In John McCarthy (ed.) *Optimality Theory in Phonology: A Reader*. Oxford: Blackwell. 495–512.

Dilley, Laura, Stephanie Shattuck-Huffnagel & Mari Ostendorf (1996). Glottalization of word-initial vowels as a function of prosodic structure. *Journal of Phonetics* **24**. 423–444.

Elias-Ulloa, José (2006). *Theoretical aspects of Panoan metrical phonology: disyllabic footing and contextual syllable weight*. PhD dissertation, Rutgers University

Féry, Caroline (2013). Focus as prosodic alignment. *Natural Language and Linguistic Theory* **31**. 683–734.

Fougeron, Cécile (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* **29**. 109–135.

Gordon, Matthew & Peter Ladefoged (2001). Phonation types: a cross–linguistic overview. *Journal of Phonetics* **29**. 383–406.

Ito, Junko & Armin Mester (1992). Weak layering and word binarity. Ms, University of California, Santa Cruz. Published in Takeru Honma, Masao Okazaki, Toshiyuki Tabata & Shin–ichi Tanaka (eds.) 2003. *A New Century of Phonology and Phonological Theory*. *A Festschrift for Professor Shosuke Haraguchi on the Occasion of His Sixtieth Birthday*. Tokyo: Kaitakusha. 26–65.

Ito, Junko & Armin Mester (2007). *Categories and projections in prosodic structure*. Handout from OCP 4, Rhodes, Greece.

Ketner, Katherine (2007). *Size restrictions in prosodic morphology*. PhD Dissertation, University of Cambridge.

Kohler, Klaus J. (1994). Glottal stops and glottalization in German. *Phonetica* **51**. 38–51.

Mester, Armin (1994). The quantitative trochee in Latin. *Natural Language and Linguistic Theory* **12**. 1–61.

Nespor, Marina & Irene Vogel (1982). Prosodic domains of external sandhi rules. In Harry van der Hulst & Norval Smith (eds.) *The Structure of Phonological Representations (Part I)*. Dordrecht: Foris. 225–255.

Pierrehumbert, Janet & David Talkin (1992). Lenition of /h/ and glottal stop. In Gerard J. Docherty & D. Robert Ladd (eds.) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. Cambridge: Cambridge University Press. 90–117.

Prince, Alan (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry* **11**. 511–562.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.

Selkirk, Elizabeth (1984). *Phonology and Syntax: the relation between sound and structure*. Cambridge: MIT Press.

Selkirk, Elisabeth (2000). The interaction of constraints on prosodic phrasing. In Merle Home (ed.) *Prosody, Theory and Experiment*. Dordrecht: Kluwer Academic Publishers. 231–261.

Selkirk, Elisabeth, Takahito Shinya & Shigeto Kawahara (2004). Phonological and phonetic effects of minor phrase length on f0 in Japanese. *Speech Prosody 2004: Proceedings of the 2nd International Conference on Speech Prosody*. Aix–en–Provence.

Shih, Chilin (1986). *The prosodic domain of tone sandhi in Chinese*. PhD Dissertation, University of California, San Diego.

Shih, Chilin (1997). Mandarin third tone sandhi and prosodic structure. In Jialing Wang & Norval Smith (eds.) *Studies in Chinese Phonology*. Walter de Gruyter. 81–123.