Interactions between prosody and morphosyntax in Fuzhou VO phrases

Aishu Chen
The Chinese University of Hong Kong

This study examines how tone sandhi domains (TSDs) are determined in Fuzhou. The data include: (i) regular verb-object phrases (VOs) where the verb takes a direct bare noun object; and (ii) non-canonical VOs where the verb takes an adverbial expression as a surface object.

Several observations are made. First, a three-way sandhi exists within every TSD. All antepenultimate syllables neutralize to low tones. A penultimate syllable's sandhi tone is dependent on the final syllable's citation tone, which remains unchanged. Second, in regular VOs, a monosyllabic verb consistently forms a single TSD with its direct bare noun object, but a disyllabic verb and its object are separated into two TSDs. Third, in non-canonical VOs, a monosyllabic verb never forms a single TSD with its adverbial object. Three questions are raised. First, what is the nature of each TSD? Second, why does the number of syllables in a verb determine the distinct TSDs formed in regular VOs? Third, how can we account for the different patterns of TSD formation in two types of VOs?

We propose that each TSD equals to a prosodic word (PrWd). OT analyses are provided to show how PrWds are derived. The analysis of regular VOs relies on the ranking of a prosodic markedness constraint FT BIN above the word-level interface correspondence constraints. The contrast between two types of VOs is explained by applying the model of Multiple Spell-Out and a cyclic interaction of morphosyntax and prosody. This approach is new in explaining the TSDs that are constrained by morphosyntax.

Keywords: tone sandhi domain, prosodic word, morphosyntax-prosody interaction, Fuzhou, VO phrases

https://doi.org/10.1075/lali.00029.che
Language and Linguistics 20:1 (2019), **46–84**
ISSN 1606-822X / E-ISSN 2309-5067

1. Introduction

This study focuses on the problem of tone sandhi domain (TSD) formation in Fuzhou post-lexical tone sandhi. Questions of current interest are: (i) What factors are involved in determining the formation of TSDs at phrase level? (ii) How do those factors interact with each other to determine the formation of TSDs?

The problem of TSD formation in Fuzhou is part of a bigger question: How shall we understand the nature of certain *domains* in terms of which some phonological or phonetic process is defined? Cross-linguistically, domain-bounded phonological phenomena are well-attested, such as the English *vowel tensing rule* (Chomsky & Halle 1968), *rhythm rule* (Liberman & Prince 1977), French *obligatory liaison rule* (Selkirk 1974), Spanish *nasal assimilation rule* (Nespor & Vogel 1982), and many others. The debate over past decades centers on how the nature of these domains should be properly defined. There are two major views regarding how to define these domains.

Some researchers maintain that the domains for phonological rule applications could be defined in terms of morphosyntactic representations. They believe that phonological rule applications can have a direct access to morphosyntactic information such as constituent structures, tree geometry, etc. This is the so-called *direct reference* approach (Kaisse 1985; Odden 1987, among others).

However, another group of researchers uphold the *indirect reference* approach. They believe that the domains for phonological processes should be defined in terms of prosodic representations, which are independent of morphosyntactic representations. This is the major claim of Prosodic Structure theory (Selkirk 1981, 1984, 1986 et seq.; Nespor & Vogel 1982, 1986; Hayes 1989; Itô & Mester 2003; Truckenbrodt 1995, among others). A widely adopted version of Prosodic Hierarchy (PH) is shown in (1).

(1) Utterance
Intonational Phrase
Phonological Phrase
Prosodic Word
Foot
Syllable
Mora

The set of hierarchically organized prosodic constituents mediates between morphosyntactic constituent structure and final phonetic output of an expression. Phonological rules take the prosodic constituent domains as the reference domains, instead of directly referring to the morphosyntactic constituent structure.

People taking the indirect reference approach maintain a close relationship between morphosyntactic constituent structure and prosodic constituent structure.

The shared view is that the prosodic constituents from the level of prosodic word above are defined in relation to morphosyntactic constituents. Different theories are posited to capture systematic correspondences between two constituent structures. Representative theories include: (i) End-based theory and Alignment constraints (M. Chen 1987; Selkirk 1986, 2000 et seq.; Selkirk & Shen 1990; McCarthy & Prince 2008; Truckenbrodt 1995); (ii) Match theory (Selkirk 2011; Elfner 2012); (iii) Relation-based theory (Nespor & Vogel 1982, 1986; Hayes 1989).

These theories diverge in how the systematic correspondence rules/constraints between the two constituent structures are formulated. First, in the edge-based theories ((i) & (ii) above), prosodic constituents are derived by referring to either one or both *edges* of morphosyntactic constituents. Second, in the relation-based theory ((iii) above), prosodic constituents are derived by considering not only the *edges* of morphosyntactic constituents, but also extra morphosyntactic information such as the head-complement relation, the direction of syntactic branching, the distinction between prefix and suffix in a morphological word, and other relational notions in syntax and morphology.

The major motivation for proposing an independent prosodic representation is the observed non-isomorphism between morphosyntactic constituents and the domains for phonological rule applications. Prosodic Structure theory provides an explanation for the non-isomorphism. The prosodic representation is proposed as an autonomous component of grammar, so there are principles governing the well-formed domination relations among prosodic constituents at different levels.[1] There are also prosodic markedness constraints which may affect the sizes or shapes of prosodic constituents.[2] When the prosodic constituent structures mapped

---

1. *Prosodic well-formedness*: the prosodic structure was once thought to be "strictly-layered" (Selkirk 1984, et al.), which means that a constituent in the hierarchy must exhaustively dominate the constituent(s) in the immediate lower level. However, later studies (Itô & Mester 2003, et al.) found that recursive and level-skipping structures are motivated by evidence from Japanese and some other languages. The "strictly-layered" requirement for well-formedness was replaced by four constraints on prosodic domination (Selkirk 2008: 466–467). They are: (i) LAYEREDNESS, which assures no constituent C dominates a constituent in the level higher than C; (ii) HEADEDNESS, which requires each constituent C to dominate at least one constituent in the lower level, except for when C is a syllable; (iii) EXHAUSTIVITY, requiring that a constituent C immediately dominates every prosodic constituent in the next level down; and (iv) NONRECURSIVITY, requiring that no constituent C could dominate another C at the same level. The first two constraints are universally held, while (iii) & (iv) can be violated in some languages.

2. *Prosodic markedness*: Examples of prosodic markedness constraints include the constraints that put restrictions on the minimal or maximal size of a prosodic constituent; the constraints that regulate the placement of prosodic prominence, etc. (cf. Selkirk 2011 for an overview of prosodic markedness constraints).

from morphosyntactic constituent structures do not satisfy prosody-internal requirements, they may be subject to further structure readjustments. As a result, the two constituent structures are not always coextensive. In Optimality Theory (OT) framework (Prince & Smolensky 1993; Kager 1999), the mapping between prosodic and morphosyntactic constituents and the structure readjustments is the result of constraint interactions, specifically the interaction of prosodic constraints with interface constraints calling for correspondences.

To sum up, domain-bounded phonological phenomena are widely attested across different languages, and the debate remains on how the nature of such domains could be properly defined. There are two major opinions. Some scholars think the domains for phonological rule application could be defined by directly referring to morphosyntactic constituent structure, but others take an indirect reference approach, admitting the role of an independent prosodic structure. For people taking the indirect reference approach, no consensus has been reached on how to properly describe the relations between morphosyntactic constituent structure and prosodic constituent structure.

The current study is concerned with Fuzhou tone sandhi at post-lexical level, which is also a domain-bounded phonological process that relies heavily on morphosyntactic information. I investigate the patterns of tone sandhi and the formation of the TSD in two types of Verb-Object phrase (VO). One is the regular VO where the verb takes a direct bare noun object. The other is the non-canonical VO where the verb takes a surface object, which is an adverbial expression (Lin 2001; 2014). I am going to demonstrate that TSDs do not necessarily correspond to any morphological or syntactic constituents. The indirect reference approach should be taken. The prosodic constituent structures which are derived based on morphosyntactic constituent structures provide the correct TSDs for Fuzhou post-lexical tone sandhi. The specific analysis is formulated in the framework of OT.

The rest of this paper is arranged as follows. More background about data sources, transcriptions, and Fuzhou tonal system are provided in § 2. The patterns of tone sandhi and the formation of TSDs are explored in § 3. It is discovered that the formation of TSDs are related to factors such as word length, word position in a phrase, and syntactic configuration. In § 4, the prosodic nature of TSDs is identified. The TSDs are argued to be prosodic words (PrWds). In § 5, I demonstrate how the two types of VO construction are parsed into PrWds under the influence of various constraints. The formation of TSDs under different circumstances are explained by PrWd-based analysis. § 6 concludes the paper, providing further evidence.

2. Language background

2.1 Data sources and transcription conventions

Fuzhou is a *Min* dialect spoken in the Northeastern part of Fujian Province in China. The broad definition of Fuzhou refers to the languages spoken in the areas around the lower reaches of the *Min* River. However, in the narrow definition, it only refers to the language spoken in the five central districts and Minhou County surrounding central Fuzhou (Li & Feng 1998). The narrow definition is taken in the current study to control the dialect-internal variations across geographical regions.

The data collected come from three sources: (i) the author's native intuition; (ii) interview data from native speakers aged 50 and above, residing in central Fuzhou city and suburban areas; (iii) published language material (Wright 1983; M. Chan 1985; Li et al. 1994; Li & Feng 1998; L. Chan 1998; Jiang-King 1995, 1996, 1999; Z. Chen 2010; Zhao 2015, among others). The materials used for interviewing are pre-constructed to control syntactic configuration, word length, etc.

Tones are transcribed using Chao's (1930) five-point scale of tone letters, where ***5*** represents the highest pitch and ***1*** represents the lowest pitch. TSDs are marked by square brackets. Sandhi tones are highlighted. Consonants and vowels are transcribed using IPA. As will be shown, vowel alternations and consonant changes sometimes occur together with tone sandhi. Alternating vowels are underlined if they occur. Consonant changes are also transcribed if they occur.

2.2 Fuzhou tonal system and diagnostics for TSDs

Li & Feng's (1998) record of Fuzhou tonal inventory is adopted here,[3] as shown in Table 1.[4]

Table 1. Fuzhou citation tones (Li & Feng 1998)

(i) Tones co-occur with tight finals	55 Yin Ping	53 Yang Ping	5̲ʔ Yang Ru	33 Shang Sheng
(ii) Tones co-occur with loose finals	212 Yin Qu	242 Yang Qu	2̲4̲ʔ Yin Ru	

3. There are many descriptive works on Fuzhou tonology. Some of them have documented slightly different tonal values. For instances, the seven citation tones are recorded as 44, 53, 5̲ʔ, 31, 213, 242, 2̲3̲ʔ in Li et al.'s (1994) system. I adopt Li & Feng's (1998) because their data are relatively updated, and their informants are citizens living in central Fuzhou city.

4. The "Yin Ping, Yang Ping, Yang Ru, Shang Sheng, Yin Qu, Yang Qu, and Yin Ru" below the tonal numbers are terms used by traditional Chinese phonologists to indicate tonal categories.

There are seven citation tones, which regularly co-occur with two types of finals,[5] either tight or loose. Besides, there are two sandhi tones: Tone35 and Tone21, which co-occur with tight finals. These two sandhi tones are not in the citation tone inventory, because they never appear together with isolated monosyllabic words in their citation forms. There is also a zero tone (T0), which appears in very few functional words, such as the aspect particle (e.g. *ko*^{0}) or the possessive particle (e.g. *ki*^{0}).

If two syllables are in the same TSD, the change of citation tone in the first syllable is often obligatory, but the final syllable remains unchanged. All logically possible combinations of two citation tones are listed below in Table 2. The leftmost column lists citation tones of the non-final syllables,[6] and the topmost row lists citation tones of the final syllables. Although there are fifty-six possible combinations of citation tones, only twenty patterns of sandhi output are got, as shown in the central part of the table. Hereafter, the sandhi patterns listed in Table 2 will be referred to as the *Disyllabic TS*.

Table 2. Disyllabic tone sandhi patterns (Li & Feng 1998)

final / non-final	55	53, 5ʔ	33	212, 242, 24ʔ
55, 212, 242, 24ʔ(i)	55–55	55–53 55–5ʔ	53–33	53–212 53–242 53–24ʔ
53, 5ʔ	55–55	33–53 33–5ʔ	33–33	21–212 21–242 21–24ʔ
33, 24ʔ(ii)	21–55	21–53 21–5ʔ	35–33	55–212 55–242 55–24ʔ

sandhi results

Except the three combinations of tone pairs "55-55", "55-53" and "55-5ʔ", all the non-final syllables within a disyllabic TSD change their citation tones, but the final syllables remain unchanged. Therefore, we can make sure that two syllables are

5. In traditional works on Chinese phonology, a syllable is divided into two parts: (i) an initial consonant, namely *initial* and (ii) the part excluding the initial consonant, namely *final*, which roughly equals to the rhyme. For example, in a syllable [tsøyʔ^{24}] 'porridge', the initial is *ts*, the final is *øyʔ*.

6. Tone24ʔ (i) and Tone24ʔ (ii) are listed as two citation tones, because they evolved from distinct tonal categories, and they behave differently in sandhi. Historically, checked tone ends with either k or ʔ. Nowadays, Fuzhou speakers in city areas no longer distinguish k and ʔ. They have merged into ʔ. In neighboring county like Gutian, speakers still distinguish k and ʔ.

within the same TSD if the preceding syllable changes its citation tone according to the *Disyllabic TS*. In cases where the citation tones of two adjacent syllables happen to form the three non-changing pairs, additional evidences are needed to judge whether the two syllables are within the same TSD or not. Besides tone sandhi, consonant changes and vowel alternations are often observed in Fuzhou continuous speech within TSDs. These two phonological changes are supplementary evidences for diagnosing TSDs.

First, consonant changes can help diagnose TSDs. When other conditions are met, the first consonant of a non-initial syllable within a TSD usually undergoes changes.[7]

(2) Consonant changes within TSDs

	Chinese	Citation tones	Sandhi tones & TSDs	Gloss
a.	照相機	kʰiɑ̯ʔ^{24} suɔ̯ŋ^{212} **ki**^{55}	[kʰia^{21} **luoŋ**^{55} **ŋi**^{55}]	'camera'
b.	衣裳	i^{55} suoŋ^{53}	[i^{55} **luoŋ**^{53}]	'clothes'
c.	修理相機	sieu^{55} li^{33} suɔ̯ŋ^{212} **ki**^{55}	[sieu^{53} li^{33}] [suo̯ŋ^{55} **ŋi**^{55}]	'to fix camera'

In (2a), the whole word forms a TSD. Consonant changes are found in the initial positions of the second syllable and the third syllable. The alveolar fricative *s* changes to an alveolar lateral *l* between two vowels, after deletion of the glottal stop coda *ʔ* in the preceding syllable. The velar stop *k* changes to a velar nasal *ŋ* when the preceding syllable ends with a velar nasal. In (2b), the two citation tones happen to form the "55-53" pair. Tonal changes are not observed in this word, however, the consonant *s* in the second syllable changes to *l*, indicating that the two syllables are within the same TSD. Otherwise, *s* would not change to *l*, like the second *s* in a separate TSD in (2c).

Vowel alternations serve as further supplementary evidence for diagnosing TSDs. Vowel alternations appear together with tone sandhi when the sandhi process involves a tone in the loose category (Table 1: ii) changing to a tone in the tight category (Table 1: i), as exemplified in (3).

(3) Vowel alternations within TSDs

	Chinese	Citation tones	Sandhi tones & TSDs	Gloss
a.	五月	ŋou^{242} ŋuoʔ^{5}	[ŋu^{55} ŋuoʔ^{5}]	'May'
b.	五，六…	ŋou^{242}, løyʔ^{5}	[ŋou^{242}] [løyʔ^{5}]	'five, six…'

7. The consonant change/lenition is basically a lexical-level phenomenon. The changes of consonants are only observed in the non-initial syllables of a TSD that is constituted by a single lexical word.

In (3a), when a tone in the loose category (Tone242) changes to a tone in the tight category (Tone55) in the first syllable, the vowel *o* is deleted, changing the rhyme *ou* to *u*. However, the same vowel in the same syllable remains the same in (3b), because the first monosyllabic word is not in a sandhi position of a TSD. The presence or absence of vowel alternations can be an additional evidence for diagnosing TSDs. However, this evidence is only useful when the tone sandhi process involves a loose tone changing to a tight tone.

In sum, rich phonological changes can be observed within TSDs. Tone sandhi sometimes are accompanied by vowel alternations and consonant changes. Through examining these phonological changes in § 3, the TSDs formed under different circumstances are identified. Further discussions on the mechanisms governing the TSDs formation are carried out in § 4 and § 5.

3. The formation of TSDs in Fuzhou VOs

3.1 Verbs with direct bare noun objects

In this section, I investigate the patterns of tone sandhi and the formation of TSDs in regular VOs. The transitive verbs take direct bare noun objects, which often have indefinite and unspecific readings.[8] Monosyllabic verbs can take the monosyllabic, disyllabic, or trisyllabic bare noun objects, and they are referred to as (1+1), (1+2), and (1+3) VOs. Similarly, when disyllabic verbs take direct bare noun objects with different lengths, they are referred to as (2+1), (2+2), and (2+3) VOs.

3.1.1 *Monosyllabic verbs with direct bare noun objects*

Examples (4) and (5) show the sandhi patterns of (1+1) VOs. Recall that Fuzhou distinguishes two types of finals: tight vs. loose, which co-occur with different types of citation tones. All possible kinds of tight-loose (T-L) combinations are explored. Some sets of data are listed below, arranged by four kinds of T-L combinations. The sandhi results of disyllabic VOs are identical to the *Disyllabic TS* patterns listed in

8. Some VO sequences are compounds, and some are idiomatic expressions. They are not included in the current investigations of post-lexical TSDs. The issue of word-phrase distinction is quite complex, and is beyond the scope of current study. Testing methods such as expansion, substitution, modification, movement, etc. are developed in Dai (1992) and others to distinguish word from phrase. Besides the syntactic ways to differentiate word from phrase, consonant changes are observed in VO compounds (e.g. 枕頭 tsieŋ^{21} **n**au^{53} 'pillow'), but not in VO phrases, even if the verb and the object are within the same TSD.

Table 2, suggesting that the monosyllabic verb uniformly forms a single TSD with its direct bare noun object.

(4) Tone sandhi in (1+1) VOs: $[\sigma]_{V.}+[\sigma]_{O.}$

	Chinese	Citation tones	Sandhi tones & TSDs	Gloss	Comb.
a.	做節	$ts\underline{ɔ}^{212}$ $ts\underline{ɑ}i\underline{ʔ}^{24}$	$[ts\underline{o}^{53}$ $ts\underline{ɑ}i\underline{ʔ}^{24}]$	'to celebrate a holiday'	L+L
b.	看店	$k^h\underline{ɑ}ŋ^{212}$ $tai\underline{ŋ}^{212}$	$[k^h\underline{a}ŋ^{53}$ $tai\underline{ŋ}^{212}]$	'to look after one's shop'	L+L
c.	賣藥	$m\underline{ɑ}^{242}$ $yo\underline{ʔ}^{5}$	$[m\underline{a}^{55}$ $yo\underline{ʔ}^{5}]$	'to sell medicine'	L+T
d.	借錢	$tsu\underline{ɔ}\underline{ʔ}^{24}$ $tsie\underline{ŋ}^{53}$	$[tsu\underline{o}^{55}$ $tsie\underline{ŋ}^{53}]$	'to borrow money'	L+T

(5) Tone sandhi in (1+1) VOs: $[\sigma]_{V.}+[\sigma]_{O.}$

a.	聽戲	$t^hia\underline{ŋ}^{55}$ $hi\underline{ɛ}^{212}$	$[t^hia\underline{ŋ}^{53}$ $hi\underline{ɛ}^{212}]$	'to listen to an opera'	T+L
b.	食卵(蛋)	$sie\underline{ʔ}^{5}$ $l\underline{ɔ}u\underline{ŋ}^{242}$	$[sie^{21}$ $l\underline{ɔ}u\underline{ŋ}^{242}]$	'to eat some eggs'	T+L
c.	搖頭	ieu^{53} t^hau^{53}	$[ieu^{33}$ $t^hau^{53}]$	'to shake one's head'	T+T
d.	鎖門	so^{33} $muo\underline{ŋ}^{53}$	$[so^{21}$ $muo\underline{ŋ}^{53}]$	'to lock door'	T+T

The tight tones and the loose tones behave asymmetrically in the non-final sandhi positions of the VOs, but they behave the same in the final positions. In (4), the preceding verbs carry citation tones of the loose category. After sandhi, the tones of the monosyllabic verbs change to the tight tones, accompanied by vowel alternations. In (5), the preceding verbs carry citation tones of the tight category. Although the citation tones of the verbs undergo value changes, there are no alternations between tight and loose categories in the sandhi processes. All tight tones remain tight after sandhi. No vowel alternations are observed. In the final positions of both (4) and (5), the citation tones of the final syllables remain unchanged, no matter whether they belong to the loose or the tight category.

Some examples of (1+2) VOs and (1+3) VOs are listed below in (6) and (7) respectively. When the bare noun objects are disyllabic or trisyllabic, the monosyllabic verbs still enter the same TSDs with their direct bare noun objects. This is revealed by the sandhi patterns.

(6) Tone sandhi in (1+2) VOs: $[\sigma]_{V.}+[\sigma\sigma]_{O.}$

	Chinese	Citation tones	Sandhi tones & TSDs	Gloss	Comb.
a.	發紅包	$hu\underline{ɑ}\underline{ʔ}^{24}$ $øy\underline{ŋ}^{53}$ pau^{55}	$[hu\underline{a}^{21}$ $øy\underline{ŋ}^{55}$ $mau^{55}]$	'to give red packets'	L TT
b.	做事情	$ts\underline{ɔ}^{212}$ $t\underline{ɑ}i^{242}$ $ki\underline{ɛ}^{212}$	$[ts\underline{o}^{21}$ $t\underline{a}i^{53}$ $i\underline{ɛ}^{212}]$	'to do things'	L LL
c.	賣麵粉	$m\underline{ɑ}^{242}$ $mi\underline{ɛ}\underline{ŋ}^{242}$ $hu\underline{ŋ}^{33}$	$[m\underline{a}^{21}$ $mi\underline{e}\underline{ŋ}^{53}$ $ŋu\underline{ŋ}^{33}]$	'to sell flour'	L LT
d.	煮元宵	tsy^{33} $ŋuo\underline{ŋ}^{53}$ $sieu^{55}$	$[tsy^{21}$ $ŋuo\underline{ŋ}^{55}$ $nieu^{55}]$	'to cook rice balls'	T TT
e.	開飯店	k^huoi^{55} $pu\underline{ɔ}\underline{ŋ}^{242}$ $tai\underline{ŋ}^{212}$	$[k^huoi^{21}$ $pu\underline{o}\underline{ŋ}^{53}$ $nai\underline{ŋ}^{212}]$	'to run a restaurant'	T LL
f.	食蹄髈	$sie\underline{ʔ}^{5}$ k^ha^{55} $pa\underline{ŋ}^{212}$	$[sie^{21}$ k^ha^{53} $\beta a\underline{ŋ}^{212}]$	'to eat pork legs'	T TL

In (1+2) VOs shown in (6), firstly, the last two syllables (i.e. the disyllabic nominal object) exhibit the pattern of *Disyllabic TS* stated above, that is, the tones in the final syllables remain unchanged, while the sandhi tones of penultimate syllables are determined by the citation tones of the final syllables. Second, regardless of their citation tones, the tones of initial monosyllabic verbs all change to a low/low falling tone (Tone 21), which belongs to the tight category.

Similar sandhi patterns are observed in (1+3) VOs shown in (7). The last two syllables (i.e. part of the nominal object) exhibit the *Disyllabic TS* pattern, that is, the sandhi tones of the penultimate syllables are determined by the citation tones of the final syllables. Regardless of their citation tones, the tones in the first two syllables (i.e. the monosyllabic verbs and the initial syllables of nominal objects) consistently change to a Tone 21.

Again, tight tones and loose tones behave asymmetrically in the non-final sandhi positions of VOs listed in (6) and (7). If a non-final syllable carries a loose citation tone, it becomes a tight tone, accompanied by vowel alternations. However, if a non-final syllable carries a tight tone in its citation form, no tight-loose alternations of tones and vowels are observed in the sandhi processes.

(7) Tone sandhi in (1+3) VOs: $[\sigma]_{V.}$+ $[\sigma\sigma\sigma]_{O.}$

	Chinese	Citation tones	Sandhi tones & TSDs	Gloss	Comb.
a.	打羽毛球	pʰɑ̲ʔ^{24} y^{33} mo^{53} kieu^{53}	[pʰa̲^{21} y^{21} mo^{33} ieu^{53}]	‘to play badminton’	**L** TTT
b.	賣清湯麵	ma̲^{242} tsʰiŋ^{55} tʰouŋ^{55}miɛŋ^{242}	[ma̲^{21} tsʰiŋ^{21} nouŋ^{53} miɛŋ^{242}]	‘to sell noodle soup’	**L** TTL
c.	借照相機	tsuɔ̲ʔ^{24} kʰiɑ̲ʔ^{24} suɔ̲ŋ^{212} ki^{55}	[tsuo̲^{21} kʰia̲^{21} luo̲ŋ^{55} ŋi^{55}]	‘to borrow a camera’	**L** LLT
d.	食芥菜飯	sieʔ^{5} kɑ̲i^{212} tsʰɑ̲i^{212} puɔŋ^{242}	[sie^{21} ka̲i^{21} ʒa̲i^{53} βuɔŋ^{242}]	‘to eat veggie rice’	**T** LLL
e.	請廚師傅	tsʰiaŋ^{33} tuo^{53} sa^{55} hɑu^{242}	[tsʰiaŋ^{21} tuo^{21} la^{53} ɑu^{242}]	‘to hire a chief’	**T** TTL
f.	騎腳踏車	kʰie^{53} kʰa^{55} taʔ^{5} tsʰia^{55}	[kʰie^{21} kʰa^{21} la^{55} ʒia^{55}]	‘to ride bike’	**T** TTT

To make an interim summary, the investigations of (1+X) VOs show that a monosyllabic verb always enters the same TSD together with its direct bare noun object, as long as the whole VO does not exceed four syllables. The tone sandhi within each TSD demonstrates two kinds of regularities. First, a *three-way distinction* of tonal behavior is observed within each TSD.[9] In particular, (i) the final syllable retains its citation tone. (ii) The sandhi tone of the penultimate syllable is determined by the citation tone of the final syllable according to the *Disyllabic TS*. (iii) The tones

9. The *three-way distinction of sandhi* is first of all observed in tone sandhi in Suzhou Chinese, which is left-dominant (Shi 2012).

of antepenultimate syllables (if any) all become a low/low falling tone, i.e. Tone 21. Second, tight tones and loose tones behave asymmetrically in the non-final sandhi positions of each TSD. Loose tones become tight tones accompanied by vowel alternations. Tight tones undergo only tone value changes without the co-occurrence of vowel alternations. However, in the final position of each TSD, syllables always retain their citation tones, no matter whether they are tight or loose.

3.1.2 *Disyllabic verbs with direct bare noun objects*

Most verbs in Fuzhou are monosyllabic, while a few verbs are disyllabic. When the disyllabic verb takes a direct bare noun object, the disyllabic verb forms an independent TSD by itself, separated from the TSD formed by the following bare noun object. This is revealed by the sandhi patterns, as shown in (8–10). The *Disyllabic TS* pattern is observed in the first two syllable domain in each VO, indicating that each disyllabic verb constitutes a single TSD. In (8), the citation tone is retained in each monosyllabic noun object. In (9), the *Disyllabic TS* is seen in each disyllabic noun object domain. In (10), the three-way sandhi pattern is present in each trisyllabic noun object. The above patterns indicate that the bare noun object in each VO forms another TSD.

The asymmetric behavior of tight tones and loose tones are again observed in the non-final sandhi positions of each TSD. When loose tones happen to appear in those non-final sandhi positions, they have to change to tight tones. However, non-final tight tones only undergo tone value changes. In the final position of each TSD, citation tones are always retained, no matter whether tight or loose.

(8) Tone sandhi in (2+1) VOs: $[\sigma\sigma]_{V.}$+ $[\sigma]_{O.}$

	Chinese	Citation tones	Sandhi tones & TSDs	Gloss	Comb.
a.	培養儂$_{(人)}$	puoi^{53} yoŋ^{33} nøyŋ^{53}	[puoi^{33} yoŋ^{33}] [nøyŋ^{53}]	‘to train somebody’	**TT** T
b.	收購菜	sieu^{55} kɑu^{212} tsʰɑi^{212}	[sieu^{53} ɑu^{212}] [tsʰɑi^{212}]	‘to buy vegetables in bulk’	**TL** L
c.	勸解儂$_{(人)}$	kʰuɔ̱ŋ^{212} ke^{33} nøyŋ^{53}	[kʰuɔ̱ŋ^{53} ŋe^{33}] [nøyŋ^{53}]	‘to comfort somebody’	**LT** T
d.	備辦乇$_{(東西)}$	pei̱^{242} paiŋ^{242} nɔʔ^{24}	[pi̱^{53} βaiŋ^{242}] [nɔʔ^{24}]	‘to prepare sth.’	**LL** L

(9) Tone sandhi in (2+2) VOs: $[\sigma\sigma]_{V.}$+ $[\sigma\sigma]_{O.}$

	Chinese	Citation tones	Sandhi tones & TSDs	Gloss	Comb.
a.	修理電器	sieu^{55} li^{33} tiɛ̱ŋ^{242} kʰei^{212}	[sieu^{53} li^{33}] [tiɛ̱ŋ^{53} ŋei^{212}]	‘to fix electronic gears’	**TT** LL
b.	交代事情	kau^{55} tɑi^{242} tɑ̱i^{242} kiɛ^{212}	[kau^{53} lɑi^{242}] [tɑ̱i^{53} iɛ^{212}]	‘to handover sth.’	**TL** LL
c.	發展教育	huɑ̱ʔ^{24} tieŋ^{33} kɑ̱u^{212} yʔ^{5}	[huɑ̱^{35} tieŋ^{33}] [kɑ̱u^{55} yʔ^{5}]	‘to develop education’	**LT** LT
d.	孝順老儂	hɑ̱u^{212} sɔuŋ^{242}lɑ̱u^{242} nøyŋ^{53}	[hɑ̱u^{53} lɔuŋ^{242}] [lɑ̱u^{55} nøyŋ^{53}]	‘to show filial piety to elders’	**LL** LT

(10) Tone sandhi in (2+3) VOs: $[\sigma\sigma]_{V.}$+ $[\sigma\sigma\sigma]_{O.}$

	Chinese	Citation tones	Sandhi tones & TSDs	Gloss	Comb.
a.	培養運動員	puoi^{53} yoŋ^{33} ouŋ^{242} touŋ^{242} ŋuoŋ^{53}	[puoi^{33} yoŋ^{33}] [uŋ^{21} nuŋ^{55} ŋuoŋ^{53}]	'to train athletes'	**TT** LLT
b.	欺負鄉下妹	k^{h}ie^{55} hou^{242} hyoŋ^{55} ŋɑ^{242} muɔi^{242}	[k^{h}ie^{53} ou^{242}] [hyoŋ^{21} ŋa^{53} muɔi^{242}]	'to bully a country girl'	**TL** TLL
c.	配合廚師傅	p^{h}uɔi^{212} haʔ^{5} tuo^{53} sa^{55} hɑu^{242}	[p^{h}uoi^{55} aʔ^{5}] [tuo^{21} la^{53} ɑu^{242}]	'to cooperate with chefs'	**LT** TTL
d.	照顧表姐妹	tsiɛu^{212} kou^{212} pieu^{33} tsia^{33} muɔi^{212}	[tsieu^{53} kou^{212}] [pieu^{21} ʒia^{55} muɔi^{212}]	'to look after one's cousins	**LL** TTL

Comparing the TSDs formed in (2+X) VOs with the TSDs formed in (1+X) VOs, the distinction is obvious. First, the monosyllabic verb consistently forms a single TSD with its direct bare noun object, but the disyllabic verb and its object are separated into two TSDs. The syllable number of the verb (i.e. being monosyllabic or disyllabic) determines the distinct TSDs formed in regular VOs.

Second, the comparison between (1+X) VOs and (2+1) VOs suggests a positional restriction on the occurrence of monosyllabic domains. The monosyllabic verbs cannot stand alone in (1+X) VOs, and they have to join the same TSDs together with the following noun objects, as in (4–7). However, the monosyllabic bare noun objects can stand alone in (2+1) VOs, as in (8). They do not join the preceding TSDs formed by the disyllabic verbs. The position where a monosyllabic word occurs determines whether it can stand alone or not. A monosyllabic word cannot stand alone in the non-final positions of a regular VO, but it can stand alone in the final position of a regular VO.

3.2 Verbs with non-canonical/adverbial objects

Although monosyllabic verbs tend to enter the same TSDs with their objects in regular VOs, such kind of tendency disappears when we consider another type of VOs where verbs take non-canonical objects. Non-canonical objects refer to those post-verbal nominal expressions that are not the logical objects of verbs, but the adverbials that modify the eventuality. Such kind of objects are called *adverbial objects* (Lin 2001; 2014). The thematic roles of the direct bare noun objects in regular VOs are normally *theme* or *patient*. Whereas the adverbial objects often bear the thematic roles such as *instrument, location, time, reason*, and so on.

(11) lists examples where the monosyllabic verbs are followed by disyllabic adverbial objects.[10] Unlike the sandhi patterns in (4–7) where the monosyllabic verbs consistently undergo tone sandhi, the monosyllabic verbs in (11) always retain their citation tones. The pattern of *Disyllabic TS* is seen in each adverbial object domain. It indicates that the monosyllabic verbs do not join the same TSDs with the following disyllabic adverbial objects. In non-canonical VOs, the verb and the adverbial object are separated into two independent TSDs.

(11) Tone sandhi in (1+2) non-canonical VOs: $[\sigma]_{V.}$+ $[\sigma\sigma]_{adv.O}$

	Chinese	Citation tones	Sandhi tones & TSDs	Gloss	Comb.
a.	做早頭	tsɔ^{212} tsai^{33} t^{h}au^{53}	[tsɔ^{212}] [tsai^{21} lau^{53}]	'to work in the morning'	**L** T T
b.	開暝晡	k^{h}uoi^{55} maŋ^{53}puo^{55}	[k^{h}uoi^{55}][maŋ^{55} muo^{55}]	'to drive in the night'	**T** T T
c.	休拜一	hyɔʔ^{24} pɑi^{212} eiʔ^{24}	[hyɔʔ^{24}] [pai^{53} eiʔ^{24}]	'to rest/close on Monday'	**L** L L
d.	開高速	k^{h}uoi^{55} ko^{55} souʔ^{24}	[k^{h}uoi^{55}] [ko^{53} louʔ^{24}]	'to drive on the highway'	**T** T L
e.	裝紙箱	tsouŋ^{55} tsai^{33} suoŋ^{55}	[tsouŋ^{55}] [tsai^{21} luoŋ^{55}]	'to pack with paper boxes'	**T** T T
f.	寫鋼筆	sia^{33} kɔuŋ^{212} peiʔ^{24}	[sia^{33}] [kouŋ^{53} meiʔ^{24}]	'to write with pen'	**T** L L
g.	打有趣	p^{h}ɑʔ^{24} ou^{242} mei^{242}	[p^{h}ɑʔ^{24}] [u^{53} mei^{242}]	'to play(cards) for fun'	**L** L L

3.3 Summary

Several findings emerge from the investigations of regular VOs and non-canonical VOs. The first finding is about the tone sandhi patterns. A three-way distinction of tone sandhi behavior is observed within each TSD with the length of two to four syllables. Only the final syllable retains its citation tone. The sandhi tone of the penultimate syllable shows a tonal dependency on the citation tone of the final syllable. The antepenultimate syllables (if any) all become Tone 21s regardless of their citation tones. At the same time, different types of tones behave asymmetrically in the sandhi positions of a TSD. Only tight tones are allowed in the non-final sandhi positions of a TSD. If loose tones occur in the non-final sandhi positions of a TSD, they will become tight tones, accompanied by vowel alternations. However, in the final position of a TSD, both tight tones and loose tones remain unchanged. The first

10. When the adverbial object construction is made up by a verb and a bare noun, other kinds of syllable combinations are not common. The disyllabic verbs are rarely followed by adverbial objects.

question arises on what kind of analysis could capture the three-way sandhi pattern in each TSD, as well as the behavior of tight and loose tones in different positions?

The patterns of TSDs formation in two types of VOs show diverse regularities, as summarized in Table 3. It is obvious that the TSDs cannot be defined by simply referring to morphological or syntactic constituents. A TSD sometimes equals to a syntactic phrase (i.a), while other times a TSD equals to a syntactic word (i.b & ii.a). No conclusion can be made on what specific types of morphological or syntactic constituents should be referred to as the TSDs. Furthermore, the syntactic configurations in (i.a) and (i.b) are the same, but the TSDs formed in them are different. It indicates that some factors other than morphosyntax would affect the formation of TSDs. Therefore, the direct reference approach does not work.

Table 3. The TSDs formed in two types of VOs (2–4 syllables)

Syntactic configurations	**TSDs formation**	
	(a) 1+X	**(b) 2+X**
(i) verb + direct bare noun object (regular VOs)	$[\text{V. Obj.}]_{\text{TSD}}$	$[\text{V.}]_{\text{TSD}}\ [\text{Obj.}]_{\text{TSD}}$
(ii) verb + adverbial object (non-canonical VOs)	$[\text{V.}]_{\text{TSD}}\ [\text{Obj.}]_{\text{TSD}}$	N/A

To figure out what precise factors would affect the formation of TSDs, the following questions need to be answered. First, the syntactic configurations in (i.a) and (i.b) are the same, why do the different lengths of the verbs (i.e. being either monosyllabic or disyllabic) lead to the distinct TSDs formed in (1+X) VOs and (2+X) VOs?

Second, in regular VOs, the position where a monosyllabic word occurs determines whether it can stand alone. The monosyllabic verb cannot stand alone in the non-final position of (1+X) VOs, so it enters the same TSD together with the following noun object. However, the monosyllabic bare noun object in the final position of (2+1) VOs can stand independently. It does not join the preceding TSD formed by the disyllabic verb. What is the reason for the asymmetric presence of monosyllabic domain?

Third, in (i.a) and (ii.a), the initial verbs are both monosyllabic. Why do the different syntactic configurations (i.e. being either a regular VO or a non-canonical VO) lead to the diverse TSDs formed in these two types of VOs?

I try to answer the above questions by taking the indirect reference approach. In § 4 and § 5, an analysis is developed in the spirit of Prosodic Structure theory. The TSDs are argued to be the prosodic constituents of designated types (i.e. PrWds). It is shown that an analysis based on PrWds provides explanations not only for the tone sandhi pattern and the behavior of two types of tones within each TSD, but also for the problem of TSDs formation under different contexts.

4. The prosodic nature of TSDs

In this section, the prosodic nature of TSDs are identified. In § 4.1, previous studies are reviewed. I briefly go over their evidences for establishing prosodic structures until the level of disyllabic foot. Based on the insights of previous works, it is argued that each TSD formed in VOs equals to a PrWd. The structures of the PrWds are outlined in § 4.2.1. I demonstrate how PrWd-internal domination and prominence relations are regulated by the prosodic markedness constraints in § 4.2.2.

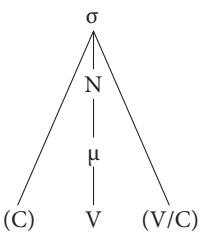
4.1 Fuzhou prosodic categories established in previous literatures

Previous studies on Fuzhou (e.g. Wright 1983; M. Chan 1985; L. Chan 1998; Jiang-King 1995, 1996, 1999, among others) have applied Prosodic Structure theory to account for various phonological phenomena, such as (i) the co-occurrence restrictions on the distributions of tones and vowels in monosyllabic words in their citation forms; (ii) the tone sandhi in a disyllabic word, and the vowel alternations with respect to the tone sandhi. The lower-level prosodic categories have been proposed, including the mora, the light/heavy syllable and the disyllabic foot.

Jiang-King's studies (1995; 1996; 1999) have established the syllable weight distinction in Fuzhou. She argues that Fuzhou distinguishes between two types of syllables: *light* vs. *heavy*. Light syllable contains only one mora, while heavy syllable contains two moras. She proposes distinct moraic structures for two types of syllables, as shown in (12), where the mora dominated by the nucleus node is the head mora.[11] The tight-loose distinction documented in descriptive literatures is argued to be the distinction between syllable weights of two types of syllables. The evidences for her argumentations on the necessity to posit the sub-syllable weight-bearing unit *mora* could be summarized in the following aspects.

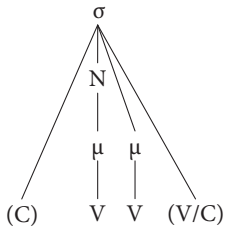
(12) Fuzhou syllable structure (Jiang-King 1996: 74)

a. light syllable (tight)



11. An on-glide occurs before the head mora sometimes.

b. heavy syllable (loose)



First, the tonal complexity in each syllable correlates with the syllable weight. Light syllables can only bear level tones or simple contour tones, whereas heavy syllables can bear complex contour tones, that is, convex or concave tones. The distributions of tones with different tonal complexities in two types of syllables indicate that the number of tone bearing units (TBUs) contained in each type of syllable is different.

Second, the number of vocalic segments that is present in each syllable correlates with the syllable weight. Specifically, (i) high vowels surface as monophthongs in light syllables, but as diphthongs in heavy syllables (e.g. i ~ ei; u ~ ou; y ~ øy). The different numbers of vocalic segments in two types of syllables indicate the quantity differences between them. (ii) Mid and low vowels appear in their tense forms in light syllables, but they appear in their lax forms in heavy syllables (e.g. a ~ ɑ; o ~ ɔ; ie ~ iɛ). Jiang-King quoted relevant phonetic studies (Fischer-Jørgensen 1990) to show that the tense-lax distinction is primarily the duration/quantity difference, but not vowel quality difference. The tense form is short, containing only one vocalic segment, while the lax form is long, containing two vocalic segments.

Third, there is a vowel harmony restriction in light syllables, but not in heavy syllables. A low-high sequence of vowels is prohibited in light syllables with consonant codas, but the low-high sequence of vowels is allowed in heavy syllables with consonant codas (e.g. eiŋ ~ aiŋ; eiʔ ~ aiʔ, where the co-occurrence of the low vowel *a* and the high vowel *i* is only possible in heavy syllables with *ʔ* or *ŋ* coda, but not in light syllables). This vowel harmony restriction indicates that the two vocalic segments are associated with one single unit in the light syllable with coda, thus the feature co-occurrence is more restrictive. Nonetheless, the feature co-occurrence is less restrictive in the heavy syllable with coda, showing that the vocalic segments are associated with more than one units.

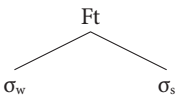
Fourth, although there are co-occurrence restrictions on tones and vowels in two types of syllables, tonal features and vowel features do not interact with each other in direct ways. The indirect interaction of tonal features and vowel features indicates that there is something mediating between tones and vowels.

Based on the facts summarized above, the prosodic category *mora* is proposed as the "anchor" for both tonal features and vocalic features. The distributions of

tones and vowels in two types of syllables, and the alternations of vowels with respect to the tone sandhi are all attributed to the different moraic structures of light and heavy syllables.[12]

Besides the light and the heavy syllables, the disyllabic right-headed foot shown in (13) has also been argued in previous studies based on the evidences from acoustic experiments, and the tone sandhi and vowel alternation patterns observed within a disyllabic word.

(13) Fuzhou right-headed foot structure
(Wright 1983; M. Chan 1985; Jiang-King 1995, 1996, 1999; L. Chan 1998, among others)



First, Wright (1983) conducted acoustic measurements, comparing the durations of two syllables in a disyllabic word with the durations of the same two syllables in isolation. Her results show that the initial syllable in a disyllabic word reduces 2/3 length compared to its citation form, whereas the final syllable reduces less than 1/3 length compared to its citation form. This result indicates that the two syllables in a disyllabic word have different degrees of prominence.[13] The initial syllable is relatively weak, so it reduces more duration. The final syllable is relatively strong, so it reduces less duration.

Second, previous studies (Wright 1983; M. Chan 1985; Jiang-King 1995 et seq.; L. Chan 1998, among others) all found that a disyllabic word constitutes a TSD. Within the TSD, the initial syllable undergoes tone sandhi, while the final syllable retains its citation tone. The change of tone in domain-initial syllable and the retention of tone in domain-final syllable suggest different degrees of prominence of the two syllables. The initial syllable is relatively weak, so tonal features in the initial syllable are not stable and subject to changes. On the other hand, the final syllable is relatively strong, being strong ensures the stabilities of tonal features.

---

12. Details on how the tonal features and the vocalic features are associated with two types of moraic structures governed by various constraints are omitted here. Interesting readers may refer to Jiang-King's work for the full analyses. Hereafter, following Jiang-King's terms, I will use *light* vs. *heavy* rather than *tight* vs. *loose* to describe different types of syllables.

13. There are acoustic studies on stress perception in Mandarin Chinese (e.g. Shen 1993), indicating that the major acoustic correlate for stress perception are duration and intensity. Duration is the most important cue in signaling stress in Chinese. Wright's (1983) results show that duration reduction in the initial position are greater than the duration reduction in the final position. Therefore, the final syllable is more prominent than the initial syllable.

The third evidence comes from the different behavior of light and heavy syllables in disyllabic tone sandhi (Jiang-King 1995 et seq.). (i) Light syllables and heavy syllables behave asymmetrically in the initial sandhi position of a disyllabic TSD. If heavy syllables appear in the initial sandhi position, the complex contour tones co-occurring with heavy syllables have to undergo simplifications, becoming level tones or simple contour tones that co-occur with light syllables (e.g. Tone242→Tone53). At the same time, the vowels of the initial heavy syllables change to their counterparts that co-occur with light syllables (e.g. p$\underline{\underline{\text{ei}}}^{242}$→p$\underline{\text{i}}^{53}$). The tonal simplifications and vowel alternations reveal that the heavy syllables in the initial position must undergo weight reduction. The weight reduction is triggered by the weak status of the initial syllable within a right-headed foot. In comparison, if light syllables appear in the initial position, no tonal simplification and vowel alternations are observed, because light syllables do not reduce weight even though they are in the weak position of a foot. (ii) Tone sandhi and vowel alternations do not take place in the final position of a TSD, no matter whether the syllable in the final position is light or heavy. This positional asymmetry again suggests that the final syllable in a disyllabic TSD is relatively strong. Both light and heavy syllables can retain their weights in the strong position of a foot, and all vocalic and tonal features are retained.

4.2 Prosodic word as the domain for tone sandhi

4.2.1 *The structure of the prosodic word*

Previous studies have established the lower-level prosodic categories in Fuzhou prosodic system, including the mora, the light / heavy syllable and the disyllabic right-headed foot. They have shown that the prosodic structures play crucial roles in explaining the disyllabic tone sandhi pattern and the behavior of light and heavy syllables. Building on their insights, I argue for a prosodic constituent above the disyllabic foot to accommodate the three-way distinction of tonal behavior, and this constituent is the PrWd. The sandhi pattern observed within each TSD is repeated below as (14).

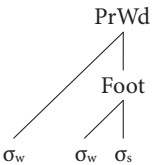
(14) Three-way distinction of tonal behavior
   a. The final syllable retains its citation tone.
   b. The penultimate syllable's sandhi tone is determined by the final syllable's citation tone according to the *Disyllabic TS*.
   c. The antepenultimate syllable(s) (if any) lose their citation tones, becoming a default low/low falling tone (Tone 21) regardless of their citation tones.

Each TSD formed in the VOs equals to a PrWd domain. The size of a PrWd can range from two to four syllables. The structures of the PrWds are shown in (15).

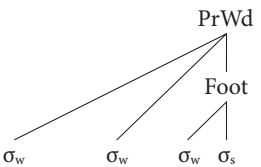
(15) Level-skipping, right-headed PrWd structures

a. Disyllabic

b. Trisyllabic



c. Tetrasyllabic



Each PrWd contains only one disyllabic right-headed foot. The foot is the head of the PrWd, and it stands on the right edge of the PrWd. In trisyllabic and tetrasyllabic PrWds, the syllables are not exhaustively dominated by the foot node in the immediate higher level. The antepenultimate syllables are directly dominated by the PrWd node, skipping the foot node.[14]

The structures of the PrWds in (15) provide an account for the three-way distinction of tonal behavior. First, the final syllable is the head of the disyllabic foot, and the foot is the head of the PrWd. Therefore, the final syllable is the strongest syllable in each PrWd. The non-final syllable(s) are either the non-headed syllable of the foot, or the unfooted syllable(s) in the PrWd. In both cases, the non-final syllables are relatively weak in the PrWd domain. The strong status ensures feature stability while the weak status leads to feature changes. That is why the syllables in the final position of a TSD always retain their citation tones, but the syllables in

14. An anonymous reviewer asks whether the proposed structures violate *Strict Layer Hypothesis*. The (15b) and (15c) are not strictly-layered, but it does not mean that they are not well-formed structures. As mentioned in footnote 1, *Strict Layer Hypothesis* has already been modified in the classic theory. Level-skipping or recursive prosodic domination is allowed.

the non-final position(s) of a TSD undergo tone sandhi. Second, within each foot domain, the sandhi tone of the penultimate syllable is determined by the final syllable's citation tone. The disyllabic foot delimits the domain for the *Disyllabic TS*. The tonal dependencies are only confined to the foot domain. Third, the unfooted weak syllable(s) dominated by the PrWd are even less prominent than the non-headed syllable of the foot. Therefore, the unfooted syllable(s) lose their citation tones, and default low tones (Tone21s) are inserted after the loss of tones.

Furthermore, the proposed structures of the PrWds in (15) could also explain the behavior of light and heavy syllables in different positions of a TSD, as summarized in Table 4.

Table 4. Behavior of light (tight) and heavy (loose) syllables in different positions of a TSD

		Antepenultimate	Penultimate	Final
Tone simplifications	*Heavy σ*	Yes	Yes	No
	Light σ	No	No	No
Vowel alternations	*Heavy σ*	Yes	Yes	No
	Light σ	No	No	No

First, the asymmetric behavior of light and heavy syllables in the penultimate or the antepenultimate position of a TSD are due to their weak status within a PrWd. The heavy syllables in the non-final weak positions need to reduce syllable weight, becoming light. The reduction of syllable weight causes tone simplifications and vowel alternations. However, the light syllables need not reduce weight even though they are in the weak positions, so the vowel alternations and tone simplifications are not present. Second, the absence of any alternations in the final position of a TSD is ensured by the strong status of the PrWd-final syllable. Both types of syllables can retain their weights, and the tonal and vocalic feature specifications are stable in the strong position.

To sum up, I have argued that the TSDs formed in VOs are equal to the PrWds domains. The structures of PrWds are outlined. Each PrWd dominates only one right-headed foot at its right edge, and one or two optional unfooted syllable(s) at its left side. The structures of the PrWds provide an account for the three-way sandhi and the behavior of light and heavy syllables within each TSD.

Besides the proposed PrWd structures in (15), alternative structures are unable to capture the three-way sandhi in each TSD. For instance, a trisyllabic unbounded foot $(\sigma_w\ \sigma_w\ \sigma_s)_{Ft}$ or a tetrasyllabic one $(\sigma_w\ \sigma_w\ \sigma_w\ \sigma_s)_{Ft}$ fails to capture the different sandhi behavior of the penultimate and antepenultimate syllables. Another alternative might be PPhs. A trisyllabic PPh in which the initial syllable links to the PPh node $\{\sigma_w\ [(\sigma_w\ \sigma_s)_{Ft}]_{PrWd}\}_{PPh}$ is consistent with the sandhi pattern, but a tetrasyllabic PPh like $\{\sigma_w\ [\sigma_w\ (\sigma_w\ \sigma_s)_{Ft}]_{PrWd}\}_{PPh}$ fails to predict the same behavior of the two antepenultimate syllables. Therefore, we choose the PrWd structures instead of others.

4.2.2 *Prosodic constraints on the structures of the prosodic words*

In § 4.2.1, the TSDs are argued to be the PrWds. The PrWds delimit the domains for the three-way sandhi. The remaining question is how the PrWds with the proposed structures are constructed in different contexts. The solution provided here follows the spirit of Prosodic Structure theory. It is believed that the formation of PrWds are governed by two types of constraints. First, the PrWds are prosodic constituents, therefore, the PrWd-internal domination and prominence relations are regulated by the prosodic markedness constraints. Second, the PrWds are defined in relation to morphosyntactic structures. The interacted effects from two sides give rise to the final outputs of PrWds. In this section, I put the focus on discussing the prosodic markedness constraints, which govern the structures of the PrWds.

Several principles should be followed to ensure that the PrWds are constructed in the proposed structures in (15). These principles are listed in (16).

(16) a. The foot is disyllabic.
b. The foot is right-headed.
c. Binary foot is constructed non-iteratively from right to left within the PrWd.
d. Every syllable should be parsed into the prosodic structure.

To realize the above principles, four prosodic markedness constraints are needed, as in (17). First, the constraint FT-BIN prohibits degenerated foot and unbounded foot from being constructed, ensuring that the foot is strictly disyllabic. Second, the constraint ALIGN (Ft, R, H(Ft), R) demands every foot to be right-headed, that is, the strongest syllable-daughter is located in final position of the foot. Third, the constraint ALIGN (Ft, R, PrWd, R) requires the right edge of a foot and the right edge of a PrWd to coincide. Within a PrWd, only one foot's right edge can be strictly aligned with the PrWd's right edge, so this constraint will be violated if there are more than one foot within a PrWd. This constraint will enforce a single foot in each PrWd if it is ranked high enough. Finally, the constraint PARSE-SYLL ensures that every syllable is parsed into a foot.

(17) FT-BIN (McCarthy & Prince 2008: 168)
"Feet must be binary under syllabic (or moraic) analysis."
ALIGN (Ft, R, H(Ft), R) (McCarthy & Prince 2008: 176)
The head of foot stands in final position in the foot.[15]
ALIGN (Ft, R, PrWd, R) (McCarthy & Prince 2008: 171)
"Every foot stands in final position in the PrWd."
PARSE-SYLL (McCarthy & Prince 2008: 168)
"All syllables must be parsed by feet."

15. This constraint is formulated following the constraint ALIGN (Ft, L, H(Ft), L) in McCarthy & Prince (2008: 176). It belongs to the family of constraints aligning the edges of prosodic constituents of different categories.

The proposed ranking of these constraints is shown in (18). As PARSE-SYLL favors feet being constructed iteratively, it must be ranked lower than ALIGN (Ft, R, PrWd, R), which militates against iterative footing. Otherwise, the output PrWd might contain more than one foot.

(18) FT-BIN, ALIGN (Ft, R, H(Ft), R), ALIGN (Ft, R, PrWd, R) >> PARSE-SYLL

Given the set of ranked constraints in (18), tableau (19) demonstrates how the optimal structures of the PrWds are constructed through the evaluations of the prosodic markedness constraints. The input is a string of two, three or four syllables without any morphosyntactic information.

First, when the input are two syllables, $Cand_a$ is selected as optimal as it satisfies all the prosodic markedness constraints. $Cand_b$ and $Cand_c$ are ruled out because they both contain degenerated foot, in violation of FT-BIN. $Cand_d$ is ruled out, because the PrWd dominates one left-headed binary foot, in violation of ALIGN (Ft, R, H(Ft), R), which favors right-headed foot. The unparsed syllable string in $Cand_e$ does not violate any of the high-ranked three constraints, but it incurs two violations of the low-ranked PARSE-SYLL, therefore, $Cand_e$ is ruled out.

Second, when the input are three syllables, $Cand_f$ is selected as optimal although it violates the low-ranked PARSE-SYLL for one time. $Cand_g$ and $Cand_h$ are ruled out due to their violations of FT-BIN. The former contains a degenerated foot, while the latter contains an unbounded foot. $Cand_i$ is ruled out because it contains a foot that does not stand in final position of the PrWd. $Cand_j$ is ruled out because it contains one left-headed foot, in violation of ALIGN (Ft, R, H(Ft), R). The unparsed $Cand_k$ is ruled out because it incurs violations of PARSE-SYLL more times than the selected $Cand_f$ does.

Third, when the input are four syllables, $Cand_l$ is selected as the optimal output although it violates PARSE-SYLL twice. $Cand_m$, $Cand_n$ and $Cand_o$ are ruled out due to their violations of FT-BIN. Each of them contains one or two feet that are not disyllabic. $Cand_p$, $Cand_q$ and $Cand_r$ are ruled out due to their violations of ALIGN (Ft, R, PrWd, R). Each of them contains a foot that does not stand in final position of the PrWd.[16] $Cand_s$ is ruled out due to the left-headed foot, which violates the constraint ALIGN (Ft, R, H(Ft), R) that militates against left-headed foot. The unparsed $Cand_t$ is ruled out because it incurs four violations of PARSE-SYLL. The selected $Cand_l$ only violates this constraint twice.

16. The violation of this alignment constraint is gradient. The time(s) of violations equal to the number of syllables intervening between the right edge of the non-final foot and the right edge of the PrWd.

(19) Tableau showing how PrWd-internal structures are regulated by the prosodic markedness constraints, supposing the input syllables are all included in a single PrWd[17]

input	output		FT-BIN	ALIGN (Ft, R, H(Ft), R)	ALIGN (Ft, R, PrWd, R)	PARSE-SYLL
σ σ	☞	a. $[(\sigma_w\ \sigma_s)]_{PW}$				
		b. $[(\sigma_s)\ (\sigma_s)]_{PW}$	*!*		*	
		c. $[\sigma_w\ (\sigma_s)]_{PW}$	*!			*
		d. $[(\sigma_s\ \sigma_w)]_{PW}$		*!		
		e. σσ				*!*
σ σ σ	☞	f. $[\sigma_w\ (\sigma_w\ \sigma_s)]_{PW}$				*
		g. $[(\sigma_s)\ (\sigma_w\ \sigma_s)]_{PW}$	*!		**	
		h. $[(\sigma_s\ \sigma_w\ \sigma_w)]_{PW}$	*!	*		
		i. $[(\sigma_w\ \sigma_s)\ \sigma_w]_{PW}$			*!	*
		j. $[\sigma_w\ (\sigma_s\ \sigma_w)]_{PW}$		*!		*
		k. σσσ				**!*
σ σ σ σ	☞	l. $[\sigma_w\ \sigma_w\ (\sigma_w\ \sigma_s)]_{PW}$				**
		m.$[(\sigma_s)\ (\sigma_s)\ (\sigma_w\ \sigma_s)]_{PW}$	*!*		*****;	
		n. $[(\sigma_s)\ (\sigma_w\ \sigma_w\ \sigma_s)]_{PW}$	*!*		***	
		o. $[(\sigma_w\ \sigma_w\ \sigma_w\ \sigma_s)]_{PW}$	*!			
		p. $[(\sigma_w\ \sigma_s)\ (\sigma_w\ \sigma_s)]_{PW}$			*!*	
		q. $[(\sigma_w\ \sigma_s)\ \sigma_w\ \sigma_w]_{PW}$			*!*	**
		r. $[\sigma_w\ (\sigma_w\ \sigma_s)\ \sigma_w]_{PW}$			*!	**
		s. $[\sigma_w\ \sigma_w\ (\sigma_s\ \sigma_w)]_{PW}$		*!		**
		t. σσσσ				***!*

To sum up, four prosodic markedness constraints are figured out, and I have demonstrated how these constraints work together to regulate the domination and prominence relations within each PrWd. (19) only shows the condition when the input syllables are all included in a single PrWd or unparsed. However, when the input are real phrases, the condition becomes different. The formation of PrWds is not only decided by the prosodic markedness constraints. It is highly related to the morphosyntactic information that are imposed on those syllables. The effects of morphosyntax are discussed in § 5.

17. The square brackets "$[\ ]_{PW}$" mark the boundaries of a PrWd, and the parentheses "( )" mark the boundaries of a foot. Different candidates show various possible PrWd-internal organizations. The solid line indicates a crucial ranking between constraints.

5. Interaction of constraints on the formation of prosodic words

5.1 The formation of prosodic words in regular VOs

In § 3.3, two questions have been raised about the TSDs formed in regular VOs. First, why does the length of the verb determine the distinct TSDs formed? Specifically, monosyllabic verb enters the same TSD together with its noun object. Disyllabic verb and its noun object are separated into two TSDs. Second, why does the position where a monosyllabic word occurs determine whether it can form an independent domain? Specifically, the non-final monosyllabic verb in a (1+X) VO cannot stand alone, but the final monosyllabic bare noun object in a (2+1) VO forms an independent domain.

As the TSDs have been argued to be the PrWds, now the questions have been transformed into the following two: (i) why does the length of the verb decide the distinct PrWds parsing? Specifically, a (1+X) VO is parsed as one PrWd, but a (2+X) VO is parsed as two PrWds. (ii) The non-final monosyllabic verb in a regular VO cannot be parsed as an independent PrWd, but the final monosyllabic noun can. What is the reason for the asymmetric tolerance of monosyllabic PrWd in different positions?

The answer lies in the interactions between the correspondence constraints and the prosodic markedness constraints discussed in § 4.2.2, especially the FT-BIN constraint. First, in the framework of Prosodic Structure theory, prosodic constituents from the level of PrWd above are defined in relation to morphosyntactic constituents by a series of correspondences constraints. Regarding the current concern about PrWds, they are defined in relation to syntactic words (WDs).[18] I adopt the MATCH constraints (Selkirk 2011), to capture the correspondences between PrWds and WDs,[19] as in (20).

---

18. In this study, the syntactic words (WDs) refer to the terminal nodes in the syntactic trees. They include: (i) individual mono-morphemic lexical items; (ii) any lexical items derived by morphological processes such as affixation or compounding.

19. There are two reasons why I adopt Match constraints, rather than the single-edge-based constraints or the relation-based constraints. First, Selkirk (2011) demonstrates that Match theory can also make correct predictions on the data that were originally taken to motivate single-edge-based theory. Furthermore, Match theory has stronger prediction power. It can explain some language data (e.g. Irish, Xitsonga) that the single-edge-based theory fails to make correct predictions. Second, there is no evidence showing that word-internal relations held between word components would affect the prosodic domain formation in the current investigations. Therefore, relational notions are not necessary to be referred to in the word-level mapping.

(20) MATCH (WD, PrWd)
The left and right edges of a WD in the input syntactic representation must correspond to the left and right edges of a PrWd in the output phonological representation.
MATCH (PrWd, WD)
The left and right edges of a PrWd in the output phonological representation must correspond to the left and right edges of a WD in the input syntactic representation.

Second, the PrWd-internal structure is regulated by the set of prosodic markedness constraints listed in (17). Among them, the constraint FT-BIN militates against degenerated foot with only one syllable. It also militates against monosyllabic PrWd, because every monosyllabic PrWd contains a degenerated foot.[20]

To account for the distinct PrWds formed in regular VOs, I propose the following ranking of the prosodic markedness constraints and the correspondence constraints, as in (21). First, FT-BIN should dominate the two MATCH constraints. Second, ALIGN (Ft, R, PrWd, R) should dominate PARSE-SYLL.

(21) FT-BIN >> MATCH (WD, PrWd), MATCH (PrWd, WD), ALIGN (Ft, R, H(Ft), R), ALIGN (Ft, R, PrWd, R) >> PARSE-SYLL

Given this ranking, a monosyllabic WD cannot faithfully map to a PrWd, because the undominant constraint FT-BIN militates against degenerated foot, and against monosyllabic PrWd. This explains why the monosyllabic verb in a regular VO cannot faithfully map to an independent PrWd. Tableaux (22–24) illustrate how PrWds are constructed in regular VOs when the verb is monosyllabic.[21]

The input in (22) is a regular (1+1) VO made up of two WDs, that are, the verb and the noun. The constraint MATCH (WD, PrWd) will be violated if one or both edge(s) of a syntactic word (WD) in the input are not aligned with the edge(s) of a PrWd in the output. MATCH (PrWd, WD) will be violated if one or both edge(s) of a PrWd in the output are not aligned with the edge(s) of a WD in the input. The output $Cand_2$ reflects a one-to-one mapping between PrWds and WDs, so the MATCH constraints are not violated. However, the two monosyllabic PrWds in $Cand_2$ incur violations of the highest-ranked FT-BIN, therefore, $Cand_2$ is ruled out.

---

20. The HEADEDNESS constraint is assumed to be inviolable in Prosodic Structure theory. A prosodic constituent must have a head. A PrWd contains at least one foot, so a monosyllabic PrWd contains a degenerated foot that is monosyllabic.

21. In the input column, the square brackets "[ ]$_{WD}$" mark syntactic word boundaries. Many candidates have been discussed in tableau (19), and those candidates are omitted in the tableaux shown in this section, which focus on the interaction between correspondence constraints and prosodic markedness constraints.

$Cand_1$ violates the two MATCH constraints for several times, because the two WDs in the input do not have their corresponding PrWds in the output. Neither does the PrWd in the output have a corresponding WD in the input. However, $Cand_1$ is still selected as the optimal output because of its satisfaction of the undominant FT-BIN.

(22) PrWd formation: regular VO, with (1+1) input
e.g. (4c) [mɑ̲^{242}]$_{V.}$ [yoʔ^{5}]$_{N.}$ → [mɑ̲^{55} yoʔ^{5}]$_{TSD}$ 'to sell medicine'

input $[\sigma]_{WD}$ $[\sigma]_{WD}$	**$Cand_1$** ☞ $[(\sigma_w\ \sigma_s)]_{PW}$	**$Cand_2$** $[(\sigma_s)]_{PW}$ $[(\sigma_s)]_{PW}$
FT-BIN		*!*
MATCH (WD, PrWd)	**	
MATCH (PrWd, WD)	*	
ALIGN (Ft, R, H(Ft), R)		
ALIGN (Ft, R, PrWd, R)		
PARSE-SYLL		

Tableau (23) shows why only one PrWd is formed in a regular (1+2) VO. The input is a monosyllabic verb followed by a disyllabic noun. $Cand_2$ reflects a one-to-one mapping between PrWds and WDs. The two WDs identically map to two PrWds. The first PrWd in $Cand_2$ contains a monosyllabic foot, thus it violates the highest FT-BIN. $Cand_2$ is ruled out due to this violation. $Cand_1$ is selected as the optimal, although it violates the MATCH constraints and the low-ranked PARSE-SYLL.

(23) PrWd formation: regular VO, with (1+2) input
e.g. (6f) [sieʔ^{5}]$_{V.}$ [kʰa^{55} pɑŋ^{212}]$_{N.}$ → [sie^{21} kʰa^{53} βɑŋ^{212}]$_{TSD}$ 'to eat pork legs'

input $[\sigma]_{WD}$ $[\sigma\sigma]_{WD}$	**$Cand_1$** ☞ $[\sigma_w\ (\sigma_w\ \sigma_s)]_{PW}$	**$Cand_2$** $[(\sigma_s)]_{PW}$ $[(\sigma_w\ \sigma_s)]_{PW}$
FT-BIN		*!
MATCH (WD, PrWd)	**	
MATCH (PrWd, WD)	*	
ALIGN (Ft, R, H(Ft), R)		
ALIGN (Ft, R, PrWd, R)		
PARSE-SYLL	*	

Tableau (24) shows why only one PrWd is formed in a regular (1+3) VO. The input is a monosyllabic verb followed by a trisyllabic noun. The output $Cand_2$, which reflects a one-to-one mapping between PrWds and WDs, is ruled out due to its violation of the undominant FT-BIN. The tetrasyllabic VO is parsed as two disyllabic PrWds

in $Cand_3$. This candidate does not incur any violations of the prosodic markedness constraints. It is ruled out because it incurs more violations of the MATCH constraints than $Cand_1$ does. Finally, $Cand_1$ wins the selection, and only one PrWd is built.

(24) PrWd formation: regular VO, with (1+3) input
e.g. (7a) $[p^h\underline{a}ɪ^{24}]_{V.}$ $[y^{33}\ mo^{53}\ kieu^{53}]_{N.}$ → $[p^h\underline{a}^{21}\ y^{21}\ mo^{33}\ ieu^{53}]_{TSD}$ 'to play badminton'

input $[\sigma]_{WD}$ $[\sigma\sigma\sigma]_{WD}$	**$Cand_1$** ☞ $[\sigma_w\ \sigma_w$ $(\sigma_w\ \sigma_s)]_{PW}$	**$Cand_2$** $[(\sigma_s)]_{PW}$ $[\sigma_w$ $(\sigma_w\ \sigma_s)]_{PW}$	**$Cand_3$** $[(\sigma_w\ \sigma_s)]_{PW}$ $[(\sigma_w\ \sigma_s)]_{PW}$
FT-BIN		*!	
MATCH (WD, PrWd)	**		**
MATCH (PrWd, WD)	*		**!
ALIGN (Ft, R, H(Ft), R)			
ALIGN (Ft, R, PrWd, R)			
PARSE-SYLL	**	*	

The above three tableaux have demonstrated why only one PrWd is constructed in a (1+X) regular VO. The highest-ranked constraint FT-BIN plays a decisive role. Due to its effect, any monosyllabic PrWd is not favored by the prosodic requirement. To avoid monosyllabic PrWd, the verb and its following noun enter the same PrWd. The ranking of FT-BIN above the two MATCH constraints leads to the mismatches between PrWds and monosyllabic verbs in (1+X) VOs. In comparison, the faithful correspondences between disyllabic verbs and PrWds do not incur violations of FT-BIN. Therefore, every disyllabic verb maps to an independent PrWd in the (2+X) regular VOs.

Then it naturally comes to the second question: why could the monosyllabic noun in the final position of a (2+1) VO form an independent PrWd? How could the phrase-final monosyllabic PrWd pass the evaluation of the undominant FT-BIN? I argue that the single syllable forms a binary foot with an empty syllable in the phrase-final/pre-pause position. As a result, the PrWd dominating this binary foot is not penalized by the highest-ranked FT-BIN. This argument is based on the insight of Duanmu's (2004; 2005) works. He proposes that a monosyllabic word in the phrase-final/pre-pause position can form a binary foot with an "empty syllable", which is also called "empty beat", "zero syllable" or "rest". The empty syllable is widely established in the literatures (e.g. Liberman 1975; Hayes 1995, among others), and I follow Duanmu in representing the empty syllable as Ø.

Tableau (25) illustrates why two PrWds are constructed in a regular (2+1) VO. The input in (25) contains a disyllabic verb followed by a monosyllabic noun. Two WDs are parsed as one PrWd in $Cand_1$. It does not violate FT-BIN, but it is ruled

out due to its violations of the two MATCH constraints. $Cand_2$ is selected as optimal because of the support of the final empty syllable. In $Cand_2$, the disyllabic verb and the monosyllabic noun identically map to two PrWds respectively, so the two MATCH constraints are satisfied. When the final syllable forms a binary foot with the empty syllable: $(\sigma_s\ \emptyset)_{Ft}$, it does not incur violation of FT-BIN. Comparing (25) with (23), the input in two cases are both trisyllabic regular VOs, but the PrWds outputs are different.

(25) PrWd formation: regular VO, with (2+1) input
e.g. (8a) [puoi^{53} yoŋ^{33}]$_{V.}$ [nøyŋ^{53}]$_{N.}$ → [puoi^{33} yoŋ^{33}]$_{TSD}$ [nøyŋ^{53}]$_{TSD}$ 'to train somebody'

Input $[\sigma\sigma]_{WD}\ [\sigma]_{WD}$	**$Cand_1$** $[\sigma_w\ (\sigma_w\ \sigma_s)]_{PW}$	**$Cand_2$** ☞ $[(\sigma_w\ \sigma_s)]_{PW}\ [(\sigma_s\ \emptyset)]_{PW}$
FT-BIN		
MATCH (WD, PrWd)	*!*	
MATCH (PrWd, WD)	*	
ALIGN (Ft, R, H(Ft), R)		
ALIGN (Ft, R, PrWd, R)		
PARSE-SYLL	*	

The PrWds formation in a (2+2) regular VO is shown in (26). $Cand_2$ is selected as optimal because it perfectly satisfies all constraints. Two WDs identically map to two PrWds in $Cand_2$. $Cand_1$ is ruled out because of its violations of the two MATCH constraints. Comparing (26) with (24), the input in two cases are both tetrasyllabic regular VOs, but the PrWds outputs are different.

(26) PrWd formation: regular VO, with (2+2) input
e.g. (9b) [kau^{55} tɑi^{242}]$_{V.}$ [tɑ̲i^{242} kiɛ^{212}]$_{N.}$ → [kau^{53} lɑi^{242}]$_{TSD}$ [tɑ̲i^{53} iɛ^{212}]$_{TSD}$ 'to hand over sth.'

input $[\sigma\sigma]_{WD.}[\sigma\sigma]_{WD}$	**$Cand_1$** $[\sigma_w\ \sigma_w\ (\sigma_w\ \sigma_s)]_{PW}$	**$Cand_2$** ☞ $[(\sigma_w\ \sigma_s)]_{PW}\ [(\sigma_w\ \sigma_s)]_{PW}$
FT-BIN		
MATCH (WD, PrWd)	*!*	
MATCH (PrWd, WD)	*	
ALIGN (Ft, R, H(Ft), R)		
ALIGN (Ft, R, PrWd, R)		
PARSE-SYLL	**	

Similarly, two PrWds are constructed in a (2+3) regular VO, as in (27). $Cand_1$ and $Cand_2$ are all ruled out due to their violations of the two MATCH constraints. $Cand_3$ reflects the identical correspondences between WDs and PrWds. It satisfies all high-ranked constraints, so it surfaces as the optimal output.

(27) PrWd formation: regular VO, with (2+3) input
e.g. (10b) [kʰie^{55} hou^{242}]$_{V.}$ [hyoŋ^{55} ŋa^{242} muɔi^{242}]$_{N.}$ → [kʰie^{53} ou^{242}]$_{TSD}$ [hyoŋ^{21} ŋa^{53} muɔi^{242}]$_{TSD}$ 'to bully a country girl'

input $[\sigma\sigma]_{WD}$ $[\sigma\sigma\sigma]_{WD}$	**$Cand_1$** $[\sigma_w\ \sigma_w\ \sigma_w\ (\sigma_w\ \sigma_s)]_{PW}$	**$Cand_2$** $[\sigma_w\ (\sigma_w\ \sigma_s)]_{PW}$ $[(\sigma_w\ \sigma_s)]_{PW}$	**$Cand_3$** ☞ $[(\sigma_w\ \sigma_s)]_{PW}$ $[\sigma_w\ (\sigma_w\ \sigma_s)]_{PW}$
FT-BIN			
MATCH (WD, PrWd)	*!*	*!*	
MATCH (PrWd, WD)	*	**	
ALIGN (Ft, R, H(Ft), R)			
ALIGN (Ft, R, PrWd, R)			
PARSE-SYLL	***	*	*

To sum up, I have illustrated how a regular VO is parsed into PrWds under the interacted effects of the prosodic markedness constraints and the interface correspondence constraints. The two questions raised at the beginning of this section are answered. First, the reason for the length of the verb to decide the PrWds formation is attributed to the ranking of FT-BIN above the two MATCH constraints. On the one hand, the MATCH constraints call for one-to-one correspondences between WDs (including V. & N.) and PrWds. On the other hand, the higher-ranked FT-BIN is against monosyllabic foot, and against monosyllabic PrWd. Due to its effect, the monosyllabic verb in a regular VO cannot identically map to a PrWd, and it joins the same PrWd together with the following noun. In comparison, the disyllabic verb in a regular VO faithfully map to a PrWd, because the PrWd formed incurs no violation of FT-BIN. Second, the empty syllable, which occurs at phrase-final position, is a remedy for the degenerated foot within a monosyllabic PrWd. The empty syllable helps a monosyllabic PrWd to pass the evaluation of FT-BIN, so that the faithful correspondence between a phrase-final/pre-pause monosyllabic WD and a PrWd could retain.

5.2 The formation of prosodic words in non-canonical VOs

The last question raised at § 3.3 is about the distinction between two types of VOs. The TSDs formation in non-canonical VOs displays a different pattern. In (11), tone sandhi is blocked between the monosyllabic verb and its adverbial object.

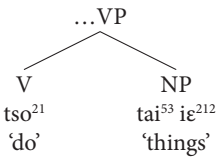
This pattern seems to contradict the PrWd-based analysis developed in § 5.1. The analysis says that a monosyllabic PrWd mapped from a monosyllabic verb will be penalized by FT-BIN, which ranks higher than the correspondence constraints. Therefore, the monosyllabic verb cannot be parsed as an independent PrWd, and it must be incorporated into the adjacent PrWd. However, in non-canonical VOs, the monosyllabic verb does not join the PrWd together with its adverbial object. How could this discrepancy be explained?

Our contention is that there are higher-level prosodic constituents PPhs, whose edges form barriers for monosyllabic incorporations. The verb and its adverbial object are parsed into different PPhs (28b), so the monosyllabic verb cannot join the same PrWd together with its adverbial object. In comparison, the verb and its logical object are in the same PPh (28a), so they could be included in the same PrWd.

(28) a. $\{ [\text{VO}]_{\text{PW}} \}_{\text{PPh}}$ vs. b. $\{ [\text{V.}]_{\text{PW}} \}_{\text{PPh}} \{ [\text{adv. O}]_{\text{PW}} \}_{\text{PPh}}$

Now the question is: what causes the different prosodic organizations in two types of VOs? Intuitively, it must be related to syntactic difference. Comparing the (1+2) regular VOs with the (1+2) non-canonical VOs, the only difference between them is their syntactic configurations. (29) shows the underlying syntactic structure of a regular VO. The monosyllabic transitive verb (V.) and the nominal phrase (NP) object are immediate constituents of a higher-level verb phrase (VP).

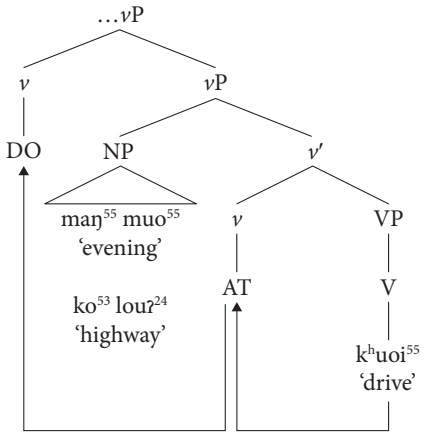
(29) Underlying syntactic structure of a regular VO



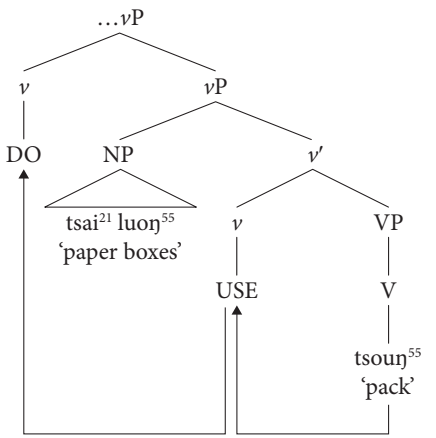
For non-canonical VOs, I adopt the syntactic analyses of Lin's (2001; 2014). (30a) shows the syntactic representation for examples like (11a–d), where the adverbial object bears the thematic role of *time* or *location*. (30b) shows the syntactic representation for examples like (11e–f), where the adverbial object bears the thematic role of *instrument*. (30c) shows the syntactic representation for examples like (11g), where the adverbial object bears the thematic role of *reason*. In each of the representation, the verbal root (lexical verb) firstly incorporates to the adverbial-introducing light verb such as AT, USE or FOR, and then it is further attracted by the subject-selecting light verb such as DO or HOLD. Finally, the verbal root conflates with two light verbs, and it appears in the position before the adverbial NP.

(30) Underlying syntactic structure of non-canonical VOs
(adopted from Lin 2001: 229–232)

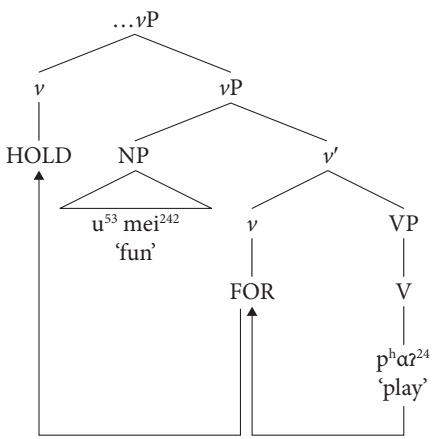
a.



b.



c.



Unlike in (29) where the verb and the NP object are sister nodes constituting a higher-level syntactic constituent immediately, the lexical verb and the adverbial object NP are not immediate constituents that directly form a higher-level syntactic constituent in the underlying structure. I hypothesize that this difference leads to the diverse patterns of prosodic structure building in the two types of VOs. However, this difference exists in the underlying structures before syntactic movements. How could prosodic structure formation refer to the underlying syntactic structure?[22]

The way to achieve this is to follow a recent thinking of syntax-prosody interface based on Phase Theory (Chomsky 2001) and Multiple Spell-Out hypothesis (Uriagereka 1999). According to the new approach, syntactic derivation proceeds in stages, and the syntactic structures are interpreted by phonology in successive cycles rather than all at once. Spell-out[23] does not happen at a single point when all syntactic computations have been finished. Instead, there are multiple applications of spell-out in a sequence of stages. Once a part of syntactic structure is ready, it gets spelled out before the full structure is constructed. This partial spell-out happens at certain points in the syntactic derivation that are called *phases*. The spell-out-domains (SODs) have been argued to correspond to prosodic constituents of various types (Kratzer & Selkirk 2007; Elfner 2012; Šurkalović 2015, among others).

There are diverse opinions on what syntactic constituents are considered as phases and when spell-out is triggered in the derivation process.[24] In the following operation, I follow Šurkalović's (2015), assuming that (i) spell-out is attempted at each syntactic merge, it is not limited to certain designated points.[25] (ii) The prosodic structure is created during each spell-out. At each phase, the elements in the SOD are prosodically parsed in certain ways. The prosodic parsing in the previous phase is stored in working memory, and is referred to in the processing of the next phase. This reference to the previous phase is achieved by introducing the *Phase-Phase Faithfulness* constraints. (iii) Spell-out proceeds in cumulative cycles, providing cumulative input to phonology. Materials that are spelled out in previous phase are still accessible in the next

---

22. An anonymous reviewer raises an important question on how reference to underlying syntactic structure can be achieved in the operation of constraint evaluation. This question helps me to think the possibility of *phase-based spell-out*, and its effects on building prosodic structure.

23. Spell-out refers to the process of linearizing the syntactic hierarchical structure and performing the lexical insertion, which retrieves the phonological underlying forms from the *Lexicon*. This is the process when syntactic structures are translated and interpreted by phonology.

24. In Chomsky's original proposal, only CP and *v*P are defined as phases. Spell-out is triggered at the point when the phase head (C or *v*) is merged into the syntactic tree. The complement of the phase head (TP for C head; VP for *v* head) gets spelled out. Once spelled out, that part of structure is impenetrable to further syntactic computation.

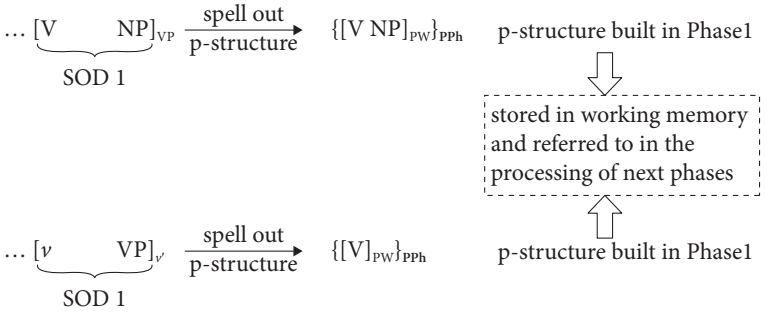
25. Epstein & Seely (2002) also argues that each application of syntactic transformation rule (merge or move) creates a phase, and triggers spell-out.

phase. The input to phonology at each phase (excluding the initial phase) includes the materials that were already spelled out and the newly merged materials.

Given the above principles in mind, let us see how different prosodic organizations in two types of VOs are derived. Simply speaking, the two types of VOs generate different prosodic structures during the initial spell-out, and the prosodic structures formed at the initial stage are retained throughout the whole derivation process. This is illustrated by (31–33).

First, in a regular VO (upper part of (31)), the monosyllabic V and the NP object are merged at the first stage, creating the VP that is evaluated immediately by phonology. At this initial phase, the monosyllabic V and the NP in the SOD are parsed as a single PrWd dominated by a PPh.[26] In a non-canonical VO (lower part of (31)), the lower *v* and the VP containing only monosyllabic V are merged at the first stage, creating the *v'*, which is sent to phonology for processing. Since the light verb *v* does not have a phonetic form, the monosyllabic V itself is parsed as a PrWd dominated by a PPh.

(31) The prosodic structures created during the first spell-out in two types of VOs



The prosodic structures created at this initial phase are stored in working memory and evaluated in the processing of next phases. The evaluation is achieved by *Phase-Phase Faithfulness* constraints, such as the one shown in (32).

(32) Phase-Phase Faithfulness constraint[27]
PHASE MAX: A prosodic constituent in phase n must have a correspondent in phase n+1 (Šurkalović 2015)

---

26. The monosyllabic V and the NP are parsed into the same PrWd, due to the reasons discussed in previous sections. If the transitive verb is disyllabic, the prosodic structure created at this stage would be a PPh dominating two PrWds.

27. This is a family of constraints, including PHASE MAX, PHASE DEP, PHASE ANCHOR, etc. For the current purpose, the MAX constraint is enough. The interactions of these constraints with other constraints (e.g. the prosodic markedness constraints) determine whether the prosodic structures built at later stages of derivation are identical to those built in the initial stage.

The prosodic structures built at the initial stage can remain identically throughout the derivation, if the Faithfulness constraint ranks high enough. This is the reason why the monosyllabic verb is parsed as an independent domain in non-canonical VOs. (33) illustrates why the initial prosodic parsing of the verb in non-canonical VO is kept unchanged in later stages of derivation. The PHASE MAX ranks higher than FT-BIN and other constraints. Parsing the verb and the adverbial object into the same PrWd/PPh would incur violation of the higher-ranked PHASE MAX. Therefore, the verb and the adverbial object are parsed separately, as in the selected candidate. The same constraint ranking would not prohibit the verb to join the same domain with its logical object, because they are parsed into the same PrWd/PPh at the initial stage of derivation.

(33) Tableau illustrating the interaction between the Phase-Phase Faithfulness and the FT-BIN

	☞ { [V.]$_{PW}$ }$_{PPh}$ { [adv. O]$_{PW}$ }$_{PPh}$	{ [V. adv. O]$_{PW}$ }$_{PPh}$
PHASE MAX		*!
FT-BIN	*	

To sum up, the different prosodic parsing of two types of VOs are attributed to their distinct underlying syntactic configurations. Under the framework of Multiple Spell-Out, the verb and its logical object are spelled out together at the initial stage, but the verb and its adverbial object are not. The higher-ranking of Phase-Phase Faithfulness constraint determines the retention of initial prosodic parsing throughout the derivation, so the verb in non-canonical VOs is kept parsed as an independent domain.

6. Conclusion

This study has developed a PrWd-based analysis to account for the varying patterns of TSDs formation in two types of VOs. Each TSD is argued to be a PrWd. It is proposed that the PrWds formation are governed by the following constraints: (i) the prosodic markedness constraints, which regulate the internal structures of PrWds; (ii) the interface correspondence constraints, which call for the faithful mappings between PrWds and syntactic words; (iii) the Phase-Phase Faithfulness constraints, which call for a retention of prosodic parsing generated at the initial stage of derivation. These constraints interact with each other, determining the final prosodic parsing.

The study is innovative in three aspects. First, it contributes empirical data by providing systematic examinations of sandhi patterns in two types of VOs with various syllable combinations. Second, it establishes PrWd structures based on

previous studies. It is the first theoretic study trying to tackle the Fuzhou TSDs formation problem by PrWds formation under constraint interactions. Third, it makes a first attempt to apply the model of Multiple Spell-Out and a cyclic interaction of syntax and prosody to account for the contrast between two types of VOs. This approach is innovative in that prosodic structures are constructed at different stages of syntactic derivation. It offers a new possibility to reconsider the long-lasting problem about tone sandhi constrained by morphosyntax. This direction is promising, as we have seen other studies such as Simpson & Wu's (2002), who takes a similar approach to explain the sandhi phenomena about Taiwanese sentence-final *Kong*. The proper phonological environment for *Kong* sandhi only appears at the intermediate syntactic derivation stage, rather than at the final stage when all syntactic movements have been finished.

Two major predictions are made based on our analysis. First, a non-phrase-final monosyllabic word and its adjacent word should join the same TSD if the two words are immediate constituents of a higher-level syntactic constituent underlyingly. This prediction is borne out by other Fuzhou phrases, such as (34a) and (34b), where the monosyllabic modifying adjective or noun joins the same TSD with the head noun, but the disyllabic adjective/noun does not, as in (34c) and (34d). The similar pattern is seen in (35a) and (35b), in which the monosyllabic adverb forms a single TSD with the following verb or adjective, but the disyllabic adverb does not, as in (35c) and (35d).

(34) Non-final monosyllabic and disyllabic Adj. / N. as modifier

a.	活蝦	uaʔ^{5} + ha^{53}	[ua^{55} ha^{53}]	alive + shrimp
b.	金戒指	kiŋ^{55} + ts^{h}ieu^{33} tsi^{33}	[kiŋ^{21} nieu^{35} ʒi^{33}]	gold + ring
c.	語文書	ŋy^{33} uŋ^{53} + tsy^{55}	[ŋy^{21} uŋ^{53}] [tsy^{55}]	Chinese + book
d.	乾淨汗衫	t^{h}ɑ̱ʔ^{24}kɑiʔ^{24} + kɑ̱ŋ^{242} louŋ^{33}	[t^{h}a̱^{53} aiʔ^{24}] [kaŋ^{53} louŋ^{33}]	clean + T-shirt

(35) Non-final monosyllabic and disyllabic Adv. as modifier

a.	乍死	tsiɑ̱^{212} + si^{33}	[tsia̱^{53} si^{33}]	just + die
b.	太厲害	k^{h}ɑ̱ʔ^{24} + lei̱^{242} hɑi^{242}	[k^{h}a̱^{21} li̱^{53} ɑi^{242}]	too + tough
c.	馬上走	ma^{33} suɔŋ^{242} + kiaŋ^{53}	[ma^{55} luɔŋ^{242}] [kiaŋ^{53}]	immediately + go
d.	確實清楚	k^{h}ɔ̱uʔ^{24} siʔ^{5} + ts^{h}iŋ^{55} ts^{h}u^{33}	[k^{h}o̱u^{21} si^{5}] [ts^{h}iŋ^{53} ʒu^{33}]	really + clear

The second prediction is about the blockage of tone sandhi across a monosyllabic word and its following word if they are not immediate constituents of a higher-level constituent underlyingly. This prediction is also supported by other phrases, such as the (36). The monosyllabic subject never enters the same TSD together with the following verb, because the prosodic parsing of the verbal predicate and the NP subject are generated at different stages of derivation.

(36) Blockage of the non-final monosyllabic incorporations in NP + V.

a.	雞叫	kie^{55} + kɔ^{212}	[kie^{55}] [kɔ^{212}]	the cock + crows
b.	汝準備錢	ny^{33} + tsuŋ^{33} pɛi^{242} + tsieŋ^{53}	[ny^{33}] [tsuŋ^{55} mɛi^{242}] [tsieŋ^{53}]	you + prepare + money

Acknowledgements

I am indebted to Jiang Ping for her insightful comments on the earlier version of this draft. This paper has benefited from extensive discussions with Prof. Jiang. I also thank two anonymous reviewers and editors from *Language & Linguistics* for their constructive comments, which help me to improve my analysis and writing. Portions of this work were presented at Hanyang University, Korea in 2015, the IACL-23 conference. Thanks also goes to the organizers and audience. All remaining errors, of course are my own.

Abbreviations

μ	mora	OT	optimality theory
σ	syllable	PPh	phonological phrase
σ_s	strong syllable	PrWd / PW	prosodic word
σ_w	weak syllable	PH	prosodic hierarchy
Adj.	adjective	SOD	spell-out-domain
Adv.	adverb	T	tight final / light syllable
adv.O	adverbial object	T-L	tight-loose
C	consonant	TP	tense phrase
Cand	candidate	TBU	tone bearing unit
CP	complementizer phrase	TSD	tone sandhi domain
Ft	foot	V	vowel
IP	intonational phrase	V.	verb
L	loose final / heavy syllable	*v*	light verb
N	syllable nucleus	VO	verb-object phrase
N.	noun	VP	verb phrase
NP	noun phrase	*v*P	light verb phrase
O.	object	WD	syntactic word

References

Chao, Yuen Ren. 1930. A system of tone letters. *Le Maître Phonétique* 45. 24–27.

Chan, Marjorie K. M. 1985. *Fuzhou phonology: A non-linear analysis of tone and stress*. Seattle: University of Washington. (Doctoral dissertation.)

Chan, Lily. 1998. *Fuzhou tone sandhi*. San Diego: University of California. (Doctoral dissertation.)

Chen, Matthew Y. 1987. The syntax of Xiamen tone sandhi. *Phonology Yearbook* 4. 109–149. https://doi.org/10.1017/S0952675700000798

Chen, Zeping. 2010. *Shijiu shiji yilai de Fuzhou fangyan* [Fuzhou dialects since 19th century]. Fuzhou: Fujian People's Press.

Chomsky, Noam & Halle, Morris. 1968. *The sound pattern of English*. Cambridge: The MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Hale, Kenneth & Kenstowicz, Michael (eds.), *Ken Hale: A life in language*, 1–52. Cambridge: The MIT Press.

Dai, Xiang-ling. 1992. *Chinese morphology and its interface with the syntax*. Columbus: The Ohio State University. (Doctoral dissertation.)

Duanmu, San, 2004. Left-headed feet and phrasal stress in Chinese. *Cahiers de Linguistique Asie Orientale* 33(1). 65–103. https://doi.org/10.3406/clao.2004.1646

Duanmu, San. 2005. The tone-syntax interface in Chinese: Some recent controversies. In Kaji, Shigeki (ed.), *Proceedings of the Symposium Cross-linguistic Studies of Tonal Phenomena, Historical Development, Tone-syntax Interface, and Descriptive Studies*, 221–254. Tokyo: Research Institute for Languages and Cultures of Asia and Africa (ILCAA), Tokyo University of Foreign Studies.

Elfner, Emily. 2012. *Syntax-prosody interactions in Irish*. Amherst: University of Massachusetts Amherst. (Doctoral dissertation.)

Epstein, S. David & Seely, T. Daniel. 2002. Rule applications as cycles in a level-free syntax. In Epstein, S. David & Seely, T. Daniel (eds.), *Derivation and explanation in the minimalist program*, 65–89. Oxford: Blackwell. https://doi.org/10.1002/9780470755662.ch4

Fischer-Jørgensen, Eli. 1990. Intrinsic F0 in tense and lax vowels with special reference to German. *Phonetica* 47. 99–140. https://doi.org/10.1159/000261858

Hayes, Bruce. 1989. The prosodic hierarchy in meter. In Kiparsky, Paul & Youmans, Gilbert (eds.), *Rhythm and meter*, 201–260. San Diego: Academic Press. https://doi.org/10.1016/B978-0-12-409340-9.50013-9

Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Itô, Junko & Mester, Armin. 2003. Weak layering and word binarity. In Honma, Takeru & Okazaki, Masao & Tabata, Toshiyuki & Tanaka, Shin-Ichi (eds.), *A new century of phonology and phonological theory: A Festschrift for Professor Shosuke Haraguchi on the occasion of his sixtieth birthday*, 26–65. Tokyo: Kaitakusha.

Jiang-King, Ping. 1995. An optimality account of tone-vowel interaction in Fuzhou. In Suzuki, Keiichiro & Elzinga, Dirk (eds.), *Proceedings of the 1995 Southwestern Workshop on Optimality Theory (SWOT)*, 89–118. Tucson: Department of Linguistics, University of Arizona.

Jiang-King, Ping. 1996. *An optimality account of tone-vowel interaction in Northern Min*. Vancouver: University of British Columbia. (Doctoral dissertation.)

Jiang-King, Ping. 1999. *Tone-vowel interaction in optimality theory* (LINCOM Studies in Theoretical Linguistics 16). Munich: LINCOM Europa.

Kaisse, Ellen M. 1985. *Connected speech: The interaction of syntax and phonology*. Orlando: Academic Press.

Kager, Rene. 1999. *Optimality theory*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511812408

Kratzer, Angelika & Selkirk, Elisabeth. 2007. Phase theory and prosodic spell-out: The case of verbs. *The Linguistic Review* 24. 93–135. https://doi.org/10.1515/TLR.2007.005

Liberman, Mark. 1975. *The intonational system of English*. Cambridge: MIT. (Doctoral dissertation.)

Liberman, Mark & Prince, Alan. 1977. On stress s and linguistic rhythm. *Linguistic Inquiry* 8. 249–336.

Lin, T.-H. Jonah. 2001. *Light verb syntax and the theory of phrase structure*. Irvine: University of California. (Doctoral dissertation.)

Lin, T.-H. Jonah. 2014. Light verbs. In Huang, C.-T. James & Li, Y.-H. Audrey & Simpson, Andrew (eds.), *The handbook of Chinese linguistics*, 73–99. Chichester: Wiley-Blackwell. https://doi.org/10.1002/9781118584552.ch4

Li, Rulong & Liang, Yuzhang & Zou, Guangchun & Chen, Zeping. 1994. *Fuzhou fangyan cidian* [Fuzhou dialect dictionary]. Fuzhou: Fujian People's Publishing House.

Li, Rong & Feng, Aizhen. 1998. *Fuzhou fangyan cidian* [Fuzhou dialect dictionary]. Nanjing: Jiangsu Education Publishing House.

McCarthy, John J. & Prince, Alan. 2008. Generalized alignment: Prosody. In McCarthy, John J. (ed.), *Optimality theory in phonology: A reader*, 167–177. https://onlinelibrary.wiley.com/doi/book/10.1002/9780470756171) (Accessed 2018-06-11)

Nespor, Marina & Vogel, Irene. 1982. Prosodic domains of external sandhi rules. In van der Hulst, Harry & Smith, Norval (eds.), *The structure of phonological representations*, vol. 1, 225–255. Dordrecht: Foris.

Nespor, Marina & Vogel, Irene. 1986. *Prosodic phonology*. Dordrecht: Foris.

Odden, David. 1987. Kimatuumbi phrasal phonology. *Phonology Yearbook* 4. 13–26. https://doi.org/10.1017/S0952675700000750

Prince, Alan & Smolensky, Paul. 1993. *Optimality theory: Constraint interaction in generative grammar*. New Brunswick: The Center for Cognitive Science, Rutgers University. (Technical Report.)

Selkirk, Elisabeth. 1974. French liaison and the $\bar{X}$ notation. *Linguistic Inquiry* 5(4). 573–590.

Selkirk, Elisabeth. 1981. On prosodic structure and its relation to syntactic structure. In Fretheim, Thorstein (ed.), *Nordic prosody II: Papers from a symposium*, 111–140. Trondheim: Tapir.

Selkirk, Elisabeth. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge: The MIT Press.

Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3. 371–405. https://doi.org/10.1017/S0952675700000695

Selkirk, Elisabeth. 2000. The interaction of constraints on prosodic phrasing. In Horne, Merle (ed.), *Prosody: Theory and experiments*, 231–261. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-015-9413-4_9

Selkirk, Elisabeth. 2008. The prosodic structure of function words. In McCarthy, John J. (ed.), *Optimality theory in phonology: A reader*, 464–482. (https://onlinelibrary.wiley.com/doi/book/10.1002/9780470756171) (Accessed 2018-06-11)

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In Goldsmith, John & Riggle, Jason & Yu, Alan C. L. (eds.), *The handbook of phonological theory*, 2nd edn., 435–484. Oxford: Blackwell. https://doi.org/10.1002/9781444343069.ch14

Selkirk, Elisabeth & Shen, Tong. 1990. Prosodic domains in Shanghai Chinese. In Inkelas, Sharon & Zec, Draga (eds.), *The phonology-syntax connection*, 313–337. Chicago: University of Chicago Press.

Shi, Xinyuan. 2012. *Tone sandhi of prosodic word in Suzhou Chinese*. Hong Kong: Chinese University of Hong Kong. (M. Phil. thesis.)

Shen, Xiaonan Susan. 1993. Relative duration as a perceptual cue to stress in Mandarin. *Language and Speech* 36. 415–433. https://doi.org/10.1177/002383099303600404

Simpson, Andrew & Wu, Zoe. 2002. Understanding cyclic spell-out. In Hirotani, Masako (ed.), *Proceedings of the North East Linguistic Society* 32, 499–518. Amherst: Graduate Linguistic Student Association (GLSA.)

Šurkalović, Dragana. 2015. *The No-Reference Hypothesis: A modular approach to the syntax-phonology interface*. Tromsø: University of Tromsø (Doctoral dissertation.)

Truckenbrodt, Hubert. 1995. *Phonological phrases: Their relation to syntax, focus, and prominence*. Cambridge: MIT. (Doctoral dissertation.)

Uriagereka, Juan. 1999. Multiple spell-out. In Epstein, S. David & Hornstein, Norbert (eds.), *Working minimalism*. Cambridge: The MIT Press.

Wright, Martha Susan. 1983. *A metrical approach to tone sandhi in Chinese dialects*. Amherst: University of Massachusetts Amherst. (Doctoral dissertation.)

Zhao, Linbin. 2015. *Fuzhouhua shiyong zidian* [Practical Fuzhou dictionary]. Shanghai: Shanghai Dictionary Press.

Author's address

Aishu Chen
Department of Linguistics and Modern Languages
The Chinese University of Hong Kong
Shatin, N.T.
Hong Kong
aishu.chen@gmail.com

Publication history

Date received: 29 June 2017
Date accepted: 27 December 2017