

The Emergence of the Binary Foot in Mandarin

Jason Brown and Shuxia Yang
University of Auckland

1 Introduction

The binary foot is a common unit utilized in prosodic morphological phenomena across languages (cf. McCarthy & Prince 1986, Hayes 1995). Formalized as a constraint, Foot-Binarity (FT-BIN) is predicted to be violable, as are other constraints in Optimality Theory. This violability implies that there exist patterns of latent foot binarity in some languages. Specifically, the existence of the constraint makes prediction that in some languages binary feet do not exist, in others binary feet play a prominent role, and yet in others binary feet will be “emergent” in some contexts.

That FT-BIN is a violable constraint is easily demonstrated in the realms of stress and morphology. For instance, degenerate and ternary feet can result in a system that demands exhaustive parsing of syllables (enforced by the constraint PARSE-SYLL) at the expense of foot binarity (PARSE-SYLL » FT-BIN); i.e. footings such as (σσσ) or (σσ)(σ) can result when not all syllables can be exhaustively parsed into binary feet. Likewise, in some languages, the prosodic size of reduplicants can emerge as a foot (all other things being equal), even though this is smaller than the base. Thus, it stands to reason that FT-BIN is violable, and as such, it also stands to reason that we should expect to find other instances of binary feet emerging in specific contexts in languages.

This paper outlines the prosodic behaviour of some compounds in Mandarin, which are demonstrated to yield output structures that must consist of binary tonal feet, and where otherwise additional obligatory syllables will undergo tonal deletion in order to satisfy this requirement. What is notable is that there appears to be only a gradient tendency for prosodic words to be equivalent to feet; there is no categorical constraint demanding it. Finally, we tackle the problem that these compounds pose: while they appear to exhibit an “emergent foot” effect, whereby an optimal unmarked structure surfaces in specific contexts, the context here is a morphosyntactically derived form, contrary to predictions of recent approaches to constraint application.

2 Incorporated Structures in Mandarin

In Mandarin, synthetic [V N] compounds, which involve the merging of a verbal root with a nominal root, exhibit a process of syllable deletion which occurs under very specific prosodic conditions. (1a) illustrates a phrasal form with a disyllabic V and a disyllabic N, and (1b) illustrates the result of compounding these categories to form [V N]:

- (1) a. ta zai shanli **xunzhao** **mogu.**
 he in mountain look.for mushroom
 ‘He looked for mushrooms in the mountain.’

*We wish to thank Michael Barlow for discussions relating to this work, as well as audience members at the Annual Meeting on Phonology 2015 in Vancouver. Much of the material in this paper is based around ideas presented in Yang (in prep), though in a much more abbreviated form. All errors are our own.

- b. ta zai shanli **xun-mo**.
 he in mountain look.for-mushroom
 'He looked for mushrooms in the mountain.'

Yang (in prep) draws similarities between this type of synthetic compounding, and the type of incorporation discussed by Hale & Keyser (2002). While we will assume that synthetic compounding involves head incorporation, nothing crucial in our analysis will hinge on it. Additional examples of this phenomenon are provided below, where (2c) illustrates the impossibility of a trisyllabic synthetic compound.

- (2) a. Ta xiu le hua-duo.
 she embroider ASP flower
 'She embroidered a flower.'
- b. Ta xiu-hua le.
 she embroider-flower ASP
 'She did some embroidering work.'
- c. *Ta xiu-hua-duo le.

Compounding is an extremely productive word-formation strategy in Mandarin, and can involve the combination of various different lexical categories (for an overview of Mandarin, see Li & Thompson 1981; for a general overview, and how Mandarin fits into a cross-linguistic context, see Aikhenvald 2007). It is, however, unclear in the literature on Chinese what distinguishes a compound from a phrase (cf. Li & Thompson 1981). Li & Thompson (1981) suggest that the constituents of compounds, as opposed to phrases, exhibit integrity and may not be separated by intervening morphemes. They also suggest that [VN] compounds do not appear with direct objects, implying an object-incorporation analysis. For these structures, aspectual particles are blocked from appearing between the verbal root and the nominal root, illustrated in (3). While phrases are free to have intervening tense/aspect/mood morphemes¹ between the verb and the noun (3a), the incorporated/compound structures are not (3b).

- (3) a. ta shuo zhe huang. (phrase)
 he say TAM lie
 'He is telling a lie.'
- b. ta shuo-huang. (synthetic compound)
 he say-lie
 'He told a lie.'

Thus, we take the placement of TAM morphemes as diagnostic for [VN] compound-hood vs. phrase-hood.

The result of a synthetic [VN] compound, as in object incorporation, is verbal. Further compounding with this incorporated stem results in nominal structures, providing further evidence that verbal compounding is bounded, and is characteristic of incorporation. For instance, consider the form: *qie cai de dao* (cut vegetable ASP knife) 'kitchen knife' (Duanmu 1997).² This is a compound form that contains an embedded compound; however, it is the inner constituent that is [VN] and thus verbal, and the larger constituent [[VN] N] is not verbal, but nominal.³ It is this boundedness that motivates an analytical focus on [VN] structures. As their prosodic output is the property in question, having a limited output (rather than an unbounded one) helps to delimit the analysis.

¹ For convenience, we have glossed the tense/aspect/mood morphemes either TAM, or simply ASP.

² Cf. the phrasal form *qie-cai-de-dao* 'a knife for cutting vegetables' (cut-vegetable-phrase marker-knife).

³ An anonymous reviewer first suggested this form to illustrate the unboundedness of compounds; however, there is reason to doubt this is even a compound form, as we take the presence of the aspect marker *de* to be diagnostic of phrases, rather than compounds.

It is the prosodic properties of the resulting verbal structure in the synthetic [VN] compounds that is of interest, and which sets this subset apart from the other types of compounds (cf. the discussion in Duanmu 2000 regarding the prosodic properties of other compound types). Regardless of the syllable count of the individual nominal and verbal roots, the resulting structure must be two syllables in length, even if this forces the deletion of a syllable (as illustrated above in examples 1 and 2). When both the verbal and nominal roots are monosyllabic, no deletion results:

(4) *ge-chang* (song-sing, ‘sing’), *shui-xi* (water-wash, ‘clean with water’) *chou-sha* (hatred-kill, ‘kill somebody because out of hatred’), *ba-ya* (pull-tooth, ‘tooth extraction’), etc.

These following illustrate all of the possible combinations of nominal and verbal roots in terms of syllable count. As (5) indicates, nouns and verbs that are monosyllabic in phrases surface unaltered when they are merged into a synthetic compound.

- (5) a. ta zai **chang** guanyu zuguo de **ge.** (phrase)
 he TAM sing about motherland ATT song
 ‘He is singing songs about his motherland.’
- b. ta zai **ge-chang** zuoguo. (synthetic compound)
 he TAM **song-sing** motherland
 ‘He is singing songs about his motherland.’

With larger structures, deletion is employed as a strategy for limiting the size of the output structure to two syllables:

- (6) a. ta zai **chang** guanyu zuguo de **ge-qu.** (phrase)
 he TAM sing about motherland ATT song
 ‘He is singing songs about his motherland.’
- b. *ta zai **ge-qu-chang** zuoguo.
 he TAM **song-sing** motherland
 ‘He is singing songs about his motherland.’
- c. ta zai **ge-chang** zuoguo. (synthetic compound)
 he TAM **song-sing** motherland
 ‘He is singing songs about his motherland.’

As can be seen from (6), (6a) is still fully grammatical with a monosyllabic verb and a disyllabic noun. (6b) is syntactically well-formed but is phonologically ruled out because of the trisyllabicity of the incorporated form. In other words, incorporation in (6b) is blocked because the noun is disyllabic. (6c), on the other hand, exhibits the same type of incorporation, but with the resulting structure being a foot.

As can be observed above, there is variation in how a noun or a verb can be expressed. For instance, in phrases, ‘song’ can be expressed as *ge*, or as *ge-qu*. Both components in the latter form independently mean ‘song’, and their use together is similar to what Inkelas & Zoll (2004) term ‘synonym constructions’, which involves the doubling of a semantic object. (7) illustrates how optional synonym doubling in phrasal forms is blocked in incorporated forms, only in the cases where the result would be a trisyllabic compound. (7a) involves a simple noun, (7b) the incorporated noun, (7c) a synonym construction, and (7d) illustrates the impossibility of incorporating the entire synonym construction.

- (7) a. ta ba le **ya.**
 he pull ASP tooth
 ‘He had a tooth extraction.’

- b. ta **ba-ya** le.
 he pull-tooth ASP
 ‘He had a tooth extraction.’
- c. ta ba le **ya-chi**.
 he pull ASP tooth
 ‘He had a tooth extraction.’
- d. *Ta **ba-ya-chi** le.
 he pull-tooth ASP

All of the logically possible prosodic patterns for forming [VN] compounds are outlined in Table 1, along with all of the attested patterns.

Table 1: Possible prosodic patterns for [V N] incorporated structures

Verb	Noun	
	σ	$\sigma \sigma$
σ	✓	✗ (repaired to $\sigma \sigma$)
$\sigma \sigma$	✗ (repaired to $\sigma \sigma$)	✗ (repaired to $\sigma \sigma$)

The generalization that emerges is that [VN] structures may result in a disyllabic foot, and no larger. The fact that a [VN] requires two constituents that are minimally monosyllabic ensures that the minimal resulting form is disyllabic, as well. We take this to be the emergence of a disyllabic foot in derived contexts. Given that there are no stray syllables allowed in this construction, the ranking required for the compounds must be FT-BIN » MAX. In the next section we will expand on the idea that the structure that results from the merging of these roots must be a disyllabic foot by claiming that it is actually a *tonal* foot that is the optimal output structure.

3 Analysis

In order to derive the foot for these compounded forms, there must be a relatively highly-ranking constraint (FTBIN) in the language. The constraint is defined in (8) as applying at the syllabic level; however, this will be amended below.

(8) FTBIN: Feet must be binary at the syllabic level.

FTBIN must dominate faithfulness to the verb and noun, respectively, which results in the deletion of excess syllables, illustrated in (9):

(9) Excess syllable deletion

/chang + gequ/	FTBIN	MAX
☞ chang-ge		*
chang-gequ	*!	

When both forms are monosyllabic FTBIN is satisfied, and MAX prevents any unnecessary deletion. When either the verb or the noun is disyllabic, however, the satisfaction of FTBIN forces deletion, and the violation of MAX.

One issue that arises is how to prevent MAX from over-applying to forms when both constituents are disyllabic.⁴ For instance, when both the verb and the noun are disyllabic, the current ranking leaves open

⁴ A related question is why only a single foot is allowed, rather than two well-formed feet. We suspect this is the product of the demand for a “perfect foot” (Ito & Mester 2015), or that [VN] compounds constitute Prosodic Words, which are not typically composed of two feet. If the former is the reason, then this raises the question of why nominal

the possibility of simply deleting both syllables from a single constituent; i.e. in order to yield a verbal form, one or both syllables from a nominal could be deleted. This strategy, as well as the strategy to achieve two well-formed feet (10b), is ruled out:

- (10) a. ta zai **yan-chang** guanyu zuguo de **ge-qu**. (phrase)
 he TAM sing about motherland ATT song
 ‘He is singing songs about his motherland.’
- b. *ta zai **ge-qu-yan-chang** zuoguo.
 he TAM song-sing motherland
 ‘He is singing songs about his motherland.’
- c. ta zai **ge-chang** zuoguo. (synthetic compound)
 he TAM song-sing motherland
 ‘He is singing songs about his motherland.’

The “hyper-deletion” strategy outlined above is prevented by REALIZEMORPH (Walker 2000, Kurisu 2001), which demands that every morpheme get a phonological expression.

- (11) REALIZEMORPH: Every morpheme has a phonological exponent in the output

The low ranking of MAX ensures that deletion will be employed to pare down the output form to meet the demands of FTBIN, but the high ranking of REALIZEMORPH ensures that deletion minimally leaves a syllable from each constituent intact.

(12) Morpheme realization

/yan-chang + ge-qu/	FTBIN	REALIZEMORPH	MAX
☞ ge-chang			*
yan-chang-ge	*!		
ge-qu		*!	

This allows for each morpheme to be recoverable from the surface form. As is also evident from the existence of monosyllabic lexical words, DEP must dominate FTBIN, otherwise every single lexical word would surface as a disyllabic foot, even if this meant augmenting an underlying monosyllabic lexical word, a state of affairs that does not obtain.

This basic analysis will be expanded on in the next section, where it will be shown that the nature of the foot in these constructions is actually tonal, and not syllabic.

4 Crowding in Tonal Feet

Previous works have argued for a tonal foot in Mandarin, with some claiming the foot is left-headed (Shih 1986, Yip 2004), and others that it is right-headed (Feng 1997). We remain agnostic as to whether feet are left- or right-headed in Mandarin, as there appear to be no constraints on the tonal structure of [VN] compounds; however, one phenomenon argues strongly for the resulting feet being tonal in nature, rather than syllable-based. This involves the appearance of classifiers in [VN] compounds, which surface between the verbal and nominal root, but with the citation tone changed to the “neutral” tone. (13) illustrates with the classifier *wan* used to modify the nominal in the compound:

- (13) a. wan (‘bowl’ classifier) Tone: 3

compounds are immune. If it is the latter, the question becomes centered on why nominal compounds would *not* constitute Prosodic Words. For the present, we leave these both as possibilities linked to open questions.

- b. he-**wan**-tang (drink-bowl-soup) ‘drink a bowl of soup’ Tones: 1 0 1

At first glance, the presence of the classifier seems to defy the general demand that incorporated structures result in disyllabic feet. Assuming that neutral-toned syllables lack a phonological tonal specification in Mandarin (Huang 2012), the existence of these structures indicates that the toneless syllable, while parsed to a tonal foot, is metrically inert in the sense that it does not count in the computation of binarity. We take the fact that these toneless syllables can intervene between syllables with specified tones to indicate that they must all belong to a single prosodic foot. The same conditions hold for other classifiers (14), as well as function words like negation (15b) and auxiliaries (15c):

- (14) a. dun (classifier) Tones: 4
 b. chi-dun-fan (eat-CL-meal) ‘have a meal’ Tones: 1-0-4
- (15) a. gan-zou (drive-leave) ‘drive somebody away’ Tones: 3 3
 b. gan-**bu**-zou (drive-NEG-leave) ‘can drive somebody away’ Tones: 3 0 3
 c. gan-**de**-zou (drive-AUX-leave) ‘can’t drive somebody away’ Tones: 3 0 3

We offer a representational solution to this problem: that the foot is tonal in nature (cf. Leben 2003). Thus, while these forms are trisyllabic, an optimal binary *tonal* foot motivates the tone deletion: since there are three morphemes, syllable deletion is not an option (as this would delete an entire morpheme, violating REALIZEMORPH); however, deleting a tone to yield a binary tonal foot achieves the optimal prosodic structure, thus FT-BIN » MAX-TONE.

Supporting evidence comes from double object constructions. Mandarin exhibits a type of “dative shift” phenomenon, whereby the two objects can invert. In addition, a preposition (marking the dative object) can appear inside of the TAM markers (which we take as diagnostic of the edge of the VP), as in (17) (tones are indicated for the bolded morphemes):

- (16) wo **song** le yi ben shu **gei ta**. (Tones: 4-3-1)
 I send TAM NUM CL book to him
 ‘I gave him a book.’
- (17) wo **song** **gei** le **ta** yi ben shu. (Tones: 4-3-1)
 I send to TAM him NUM CL book
 ‘I gave him a book.’

This is in contrast to structures where a pronominal form can also incorporate into the verbal complex. In these cases, the restriction is on pronominal elements, and when this happens, the pronominal form becomes toneless, presumably due to the constraint on tonal feet proposed above:

- (18) wo **song** **gei ta** le yi ben shu. (Tones: 4-3-0)
 I send to him TAM NUM CL book
 ‘I gave him a book.’

Only pronominals are allowed in these structures, and the tonal reduction in these cases is paralleled by the cases of classifiers and other functional elements. These examples illustrate the point that the conditions on binary feet are widespread through the language, and not limited to synthetic compounds. Multiple types of these structures are analyzed by Yang (in prep).

Finally, it should be noted that the notion of a tonal foot is not novel for Chinese languages. This is also the structure that has been argued to be at the heart of compounds in other Chinese languages (Duanmu 1997, 2000). This discussion of Mandarin extends that notion. What remains is a comprehensive overview of synthetic compounds in all of the Chinese languages.

5 Implications Surrounding FT-BIN

In recent work on English and Navajo, Martin (2011) has claimed that phonological constraints can have a categorical effect within lexical items that is expressed as a gradient effect in larger, derived contexts such as compounds (cf. also Mohanan 1993 for the idea that constraints apply more stringently to smaller domains). Martin discusses segmental constraints that are active in English and Navajo words, but which only exhibit gradient effects in compounds. Martin's analysis involves viewing the word-level constraints as applying less strongly in derived contexts, such as compounds.

The problem raised by the Mandarin compound pattern is that the binary foot effect holds in [VN] compounds, but FT-BIN is freely violable in lexical forms. For example, Zhou (2004) found that the *Xinhua New Word Dictionary* includes 2168 words, among which 1204 are disyllabic, 324 are trisyllabic, and 449 are quadrisyllabic, and that *The Contemporary Chinese Dictionary* has 58,481 words, among which 39,548 are disyllabic, 4,828 are trisyllabic, and 4,798 are quadrisyllabic. The implication is that within lexical items, FT-BIN is violable - there are numerous lexical forms to attest to this (including several of the monosyllabic verbs found in the examples above). However, the force of FT-BIN clearly emerges as a statistical preference over the lexicon. This gradiently violated lexical constraint is satisfied categorically in these particular derived contexts. We thus claim, based on the above patterns, that phonological constraints may in some cases apply more stringently to larger or derived domains than to lexical items (the *opposite* of Martin's claims for English). We liken the Mandarin patterns to other phenomena such as reduplication, where derived environments provide the context for emergent structure. Like Martin's findings, though, the bias in the pattern informs the categorical constraint; in this case, the gradient lexical pattern is manifested as an absolute prohibition in derived verbal structures.

The use of the term "emergent" implies an emergence of the unmarked type of effect, and is entirely consistent with the emergence of prosodic morphology in derived environments, such as reduplication (cf. McCarthy & Prince 1994). Under this view, the output of a morphological operation results in unmarked structures, where the constraint prohibiting these structures is normally violable in the language. Under TETU, the lack of a correspondence relation between a structure in question (e.g. an epenthetic segment or a reduplicant) and the input is what allows a markedness constraint to make decisions that favor unmarked structures. However, in the present case, the correspondence relations are not entirely clear. Furthermore, the statistical tendency towards binary feet would remain unexplained. Thus, we leave these avenues for future research.

6 Conclusion

While much is known about the phonology of Mandarin, there is still much to be discovered about the phonology-syntax interactions, and the prosodic morphology that is associated with those operations. We have argued that there is a tonal foot in the language which is responsible for the prosodic shapes that are found in synthetic compounds. This investigation, however, raises several important questions. These include: To what extent do phonological processes mirror lexical patterns? And in what contexts are emergent structures like feet allowed to surface? Reduplication is one, and here we have shown that compounds or incorporated forms (including in double object constructions) are also valid structures for emergent prosody.

7 References

- Aikhenvald, Alexandra Y. 2007. Typological distinctions in word-formation. In Timothy Shopen (ed.), *Language typology and syntactic description: Grammatical categories and the lexicon, 2nd edition*. Cambridge: Cambridge University Press, pp. 1-65.
- Duanmu, San. 1997. Phonologically motivated word order movement: evidence from Chinese compounds. *Studies in Linguistic Sciences* 27:49-73.
- Duanmu, San. 2000. *The phonology of Standard Chinese*. Oxford: Oxford University Press.
- Feng, Shengli. 1997. *Interaction between morphology, syntax and prosody in Chinese*. Beijing: Peking University Press.
- Hale, Ken & Samuel Jay Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.

- Hayes, Bruce. 1995. *Metrical stress theory*. Chicago: University of Chicago Press.
- Huang, Karen. 2012. A study of neutral-tone syllables in Taiwan Mandarin. Ph.D. dissertation, University of Hawai'i at Mānoa.
- Inkelas, Sharon & Cheryl Zoll. 2004. *Reduplication: Doubling in morphology*. Cambridge: Cambridge University Press.
- Ito, Junko & Armin Mester. 2015. The perfect prosodic word in Danish. *Nordic Journal of Linguistics* 38:5–36.
- Kurusu, Kazutaka. 2001. The phonology of morpheme realization. Ph.D. thesis, University of California, Santa Cruz.
- Li, Charles N. & Sandra A. Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Leben, William R. 2003. Tonal feet as tonal domains. In John M. Mugane (ed.), *Linguistic typology and representation of African languages*. Trenton, NJ: Africa World Press, Inc., pp. 129-138.
- Martin, Andrew. 2011. Grammars leak: Modeling how phonotactic generalizations interact within the grammar. *Language* 87:751-770.
- McCarthy, John J. & Alan S. Prince. 1986. Prosodic morphology. Ms, University of Massachusetts, Amherst & Brandeis University.
- McCarthy, John J. & Alan S. Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. In M. González (ed.), *Proceedings of the North East Linguistic Society* 24. Amherst, MA: GLSA, pp. 333-379.
- Mohanan, K. P. 1993. Fields of attraction in phonology. In John Goldsmith (ed.), *The last phonological rule: Reflections on constraints and derivations*. Chicago: University of Chicago Press, pp. 61-116.
- Shih, Chilin. 1986. The prosodic domain of tone sandhi in Chinese. Ph.D. dissertation, UCSD.
- Walker, Rachel. 2000. *Nasalization, neutral segments, and opacity effects*. New York: Garland.
- Yang, Shuxia. In prep. The phonology of incorporation in Mandarin. Ph.D. dissertation, University of Auckland.
- Yip, Moira. 1980. The tonal phonology of Chinese. Ph.D. dissertation, MIT.
- Yip, Moira. 2003. Feet and tonal reduction at the word and phrase level in Chinese. In J. J. McCarthy (ed.), *Optimality Theory in phonology: A reader*. Malden, MA: Blackwell Publishing, pp. 228-245.
- Zhou, J. 2004. *On Chinese word structure*. Shanghai Lexicographical Publishing House.