Yang, Yifan. 2024. Floating mora affixation in Huozhou diminutive subtraction. *Glossa: a journal of general linguistics* 9(1). pp. 1–41. DOI: https://doi.org/10.16995/glossa.15130

Open Library of Humanities

Floating mora affixation in Huozhou diminutive subtraction

Yifan Yang, Shanghai Jiao Tong University, CN, yifanyang@sjtu.edu.cn

This paper argues for an item-and-arrangement approach by demonstrating that the diminutive form in Huozhou Chinese results from the affixation of a floating mora. Diminutives in Huozhou Chinese are formed by deleting the final non-vocalic segment of a syllable, along with some subsegmental changes (e.g., [pʰɑŋ^{35}] 'plate', [pʰaː^{35}] 'plate.DIM'). This paper provides a detailed description of these patterns and proposes that the underlying representation of the diminutive morpheme is a floating mora, with subtraction serving as a repair strategy for this floating mora affix. This paper makes two main contributions. First, it introduces new data on subtractive morphology. Second, it provides a formal analysis and further supports this proposal from a typological perspective, thereby supporting the item-and-arrangement approach to morphology. Overall, by combining a case study of Huozhou Chinese with a typological discussion, this analysis shows that nonconcatenative morphology can be interpreted as additive, leading to a more economical grammar and more restrictive predictions.

1 Introduction

Most morphological constructions in languages are formed through the sequential concatenation of phonological material paired with meaning, such as the English plural [dɔg-z]. However, other types of morphological construction are achieved by phonological processes such as copying, deleting, or feature changing. In Tohono O'odham, for instance, perfective verbs are formed by deleting the final segment of their imperfective counterparts (e.g., [hiːnk] 'bark.IPFV' and [hiːn] 'bark.PFV') (Anderson 1992; cf. Stonham 1994). Morphological constructions that exhibit rule-like operations are called nonconcatenative morphology, and common examples include reduplication, subtractive morphology, morphological lengthening, and morphological mutation.

These two types of morphology, concatenative and nonconcatenative, are crucial to a debate on two approaches to morphology: item-and-arrangement and item-and-process (Hockett 1954). The item-and-arrangement approach involves identifying morphemes (*items*) as fundamental units and understanding how these morphemes are combined (*arranged*) to form words. Since most morphological constructions are concatenative, they can easily be treated by item-and-arrangement models, where the phonological form of each morpheme is arranged in an additive fashion (e.g., Selkirk 1982; Lieber 1992; Halle & Marantz 1993, a.o.). However, nonconcatenative morphology, such as the Tohono O'odham example, challenges this approach. In contrast, the item-and-process approach assumes that complex words are formed through operations on simpler ones. Rule-like nonconcatenative morphology has been interpreted as evidence for item-and-process models (e.g., Anderson 1992; Kurisu 2020). In general, item-and-arrangement models are argued to be more restrictive, while item-and-process models can better account for nonconcatenative morphology (Anderson 1992: 63ff).

Extensive research over the past decades has suggested that many instances of nonconcatenative morphology can be reanalyzed as the concatenation of phonological material, especially with the development of Autosegmental Phonology (Goldsmith 1976), Moraic Theory (Hyman 1984; 1985; Hayes 1989), and Prosodic Morphology (McCarthy & Prince 1986/1996). A recently thriving theory of morphology and phonology, Generalized Nonlinear Affixation (GNA, Bermúdez-Otero 2012), seeks to reanalyze all types of nonconcatenative morphology as the affixation of nonlinear phonological elements, such as floating features, tones, and prosodic nodes (mora, syllable, foot, etc.). Some phenomena that have been reanalyzed as concatenation and considered as arguments for GNA include morphological lengthening/gemination (e.g., Davis and Ueda 2002; 2006), reduplication (e.g., Saba Kirchner 2010; 2013; Bye and Svenonius 2012; Zimmermann 2013), and morphological mutation (e.g., Akinlabi 1996; Wolf 2007; Paschen 2018).

Subtractive morphology, however, poses a significant challenge to the item-and-arrangement approach, as it involves the shortening of a morphological base. Although subtraction appears opposite to additive morphology in nature, some recent studies advocate for an item-and-arrangement analysis of subtractive morphology (e.g., Trommer 2011; Bye & Svenonius 2012;

Trommer and Zimmermann 2014; Zimmermann 2017; Köhnlein 2018). In these studies, subtractive morphology has been argued to result from the affixation of an underspecified root node (Bye & Svenonius 2012), a floating mora (e.g., Trommer and Zimmermann 2014), or a foot template (Köhnlein 2018). All of these analyses are considered arguments in favor of GNA, an item-and-arrangement model.

This paper aims to investigate the subtractive patterns of diminutive formation in Huozhou Chinese, emphasizing its phonological aspects, which provide further evidence for an item-and-arrangement approach, particularly GNA. In Huozhou Chinese, diminutives can be formed by subtraction, exemplified as follows (the superscript digits indicate tones, where "1" indicates the lowest pitch while "5" indicates the highest).

(1) Diminutive subtraction in Huozhou Chinese

	Non-diminutive	Diminutive	Gloss
a.	[sa**j**^{11}]	[saː^{53}]	'sieve.DIM'
b.	[ŋɑ**w**^{53}]	[ŋoː^{53}]	'quilted coat.DIM'
c.	[p^{h}ɑ**ŋ**^{35}]	[p^{h}aː^{35}]	'plate.DIM'

As shown in (1), the diminutives in Huozhou Chinese can be formed by deleting the last non-vocalic segment, along with featural changes. In this paper, drawing upon Lin's (1993) approach to similar morphophonological patterns in Chinese dialects (see §4 for details), I propose that the phonological form of the Huozhou diminutive morpheme is a floating mora. A mora (μ) is a prosodic unit that represents weight or duration, and it helps organize segments into larger prosodic structures, such as syllables and feet. A floating mora is a mora that has no associated segments and has not been integrated into a prosodic structure. In this way, the phonological form of a diminutive noun can be viewed as the concatenation of a root and a floating mora affix (e.g., $/\mathrm{sa}^{\mu}\mathrm{j}^{\mu} + \mu/$), while the subtracted forms in (1) result from the repair of the floating mora affix by the phonological grammar. Further, additional arguments for floating mora affixation come from a typological perspective. In nearby related dialects, diminutives can be formed by lengthening, segment epenthesis, or merger, all of which can be analyzed as the result of floating mora affixation. In a broad sense, this proposal is in line with Trommer (2011), Trommer & Zimmermann (2014), Zimmermann (2017), and Köhnlein (2018), where subtractive morphology is analyzed as prosodic node affixation (of a mora in particular) and nonconcatenative morphology is viewed as an epiphenomenon of phonology.

The goals of this paper are twofold. First, Huozhou Chinese remains under-studied in generative linguistics, though the special characteristics of diminutive formation have long been noted in descriptive works (Satoh & Feng 1989; X. Tian 1992; Hou & Wen 1993; Shen et al. 2010; Feng & Zhao 2014). To address this, this paper aims to conduct a thorough formal phonological analysis of the observed patterns within the framework of Optimality Theory (Prince & Smolensky 1993/2004), based on the author's fieldwork data. Second, the analysis shows that Huozhou diminutive subtraction can be analyzed as floating mora affixation, which provides arguments for the item-and-arrangement approach, particularly the GNA model, and leads to twofold parsimony. On the one hand, the moraic analysis of Huozhou Chinese is more restrictive and parsimonious compared to other theories in OT that are akin to the item-and-process approach (e.g., Alderete 2001; Kurisu 2001). On the other hand, this paper discusses similar patterns in nearby dialects (lengthening, epenthesis, and merger) and demonstrates that they can be analyzed as different surface realizations of floating mora affixation. This provides additional arguments for GNA from a typological perspective since dialects of the same language demonstrate both additive and subtractive effects of floating mora affixation.

The rest of the paper is organized as follows. §2 presents the data; §3 lays out the theoretical proposal and provides a formal analysis of the observed patterns; §4 examines the diminutive forms in nearby dialects; §5 compares the current proposal with other item-and-process alternatives in OT, and §6 concludes the paper. In the following discussion, "mora affixation" refers specifically to the affixation of a floating mora.

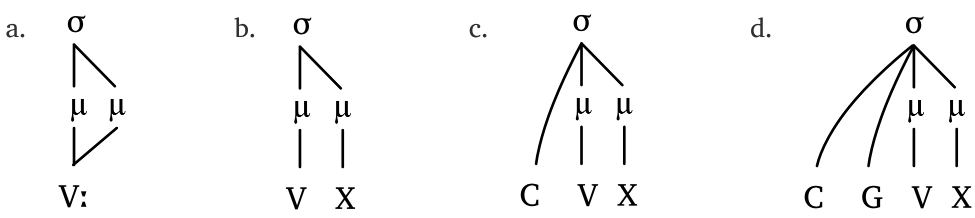
2 Diminutive formation in Huozhou Chinese

2.1 Huozhou Chinese and its phonology

Houzhou is located in the Fenhe River valley, southern Shanxi Province. Huozhou Chinese is classified as a variety of Zhongyuan Mandarin (Shen 2008), and there are four sub-varieties: West Huozhou, East Huozhou, North Huozhou, and Wangzhuang (Feng & Zhao 2014: 10).

Huozhou Chinese shares the same syllable structure as Standard Chinese. The maximal syllable template is CGVX, in which G represents a pre-nuclear glide ([j], [w], or [ɥ]), and X can be either a post-nuclear glide ([j] or [w]) or a velar nasal ([ŋ]). Although there have been various proposals regarding the internal structure of Chinese syllables (e.g., Cheng 1966; Lin 1989; Bao 1990; 1996; Duanmu 1990; 2007; van de Weijer & Zhang 2008), it is the moraic representation of Chinese syllables that is relevant to the discussion in this paper, as illustrated in (2) (after Zhang 2023a: 4).

(2) Moraic representations of Huozhou Chinese syllables (tones are omitted)



In Standard Chinese, syllables that carry a full lexical tone are treated as bimoraic (Duanmu 1990; 1993; 1994; 2007; Wu and Kenstowicz 2015). I assume that Huozhou Chinese shares the same property. Therefore, the nuclear vowel in an open syllable is represented as bimoraic on the surface to bear the lexical tone, as in (2a). Meanwhile, only elements in the rime domain (V or VX) can be moraic; namely, moraicity does not extend to the prenuclear glide (2b–d).

Since only syllables that end with [ŋ], [j], or [w] are closely related to the observed patterns, the possible (G)VX combinations in Huozhou are listed in (3), with their corresponding underlying forms in parentheses.

(3) (G)VX combinations of Huozhou Chinese

a. ŋ-ending				b. j-ending			c. w-ending	
ɑŋ (/aŋ/)	jɑŋ (/iaŋ/)	wɑŋ (/uaŋ/)	ɥɑŋ (/yaŋ/)	aj (/ai/)	jaj (/iai/)	waj (/uai/)	ɑw (/au/)	jɑw (/iau/)
əŋ (/əŋ/)	iŋ (/iŋ/)	uŋ (/uŋ/)	yŋ (/yŋ/)	ej (/əi/)		wej (/uəi/)	ow (/əu/)	jow (/iəu/)

Regarding the underlying representations, both onglides and offglides are derived from their corresponding vowels (/i/ → [j], /y/ → [ɥ], and /u/ → [w]). The low vowel /a/ can be either front [a] or back [ɑ], depending on the influence of the postvocalic element, as in [jɑŋ] (*[jaŋ]) and [waj] (*[wɑj]). Similarly, /ə/ surfaces as [o], [e], and [ə] when followed by [w], [j], and [ŋ], respectively, which is a common phenomenon in Chinese syllables (Duanmu 2007; Lin 2015, among others).[1] The feature specifications of major surface vowels are given in (4) to facilitate the discussion.

(4) Feature specifications of major surface vowels

	i	y	u	e	ə	o	a	ɑ
[high]	+	+	+	–	–	–	–	–
[low]	–	–	–	–	–	–	+	+
[back]	–	–	+	–	+	+	–	+
[round]	–	+	+	–	–	+	–	–

Finally, there are five lexical tones in Huozhou Chinese: low-level, rising, mid-level, falling, and high-level. In this paper, I transcribe the tones with Chao's (1968) digit system as "11", "35", "33", "53", and "55", where "1" represents the lowest pitch while "5" represents the highest.

[1] The prenuclear glide can also affect the quality of the nuclear vowel. In (C)GV: syllables, /a/ and /ə/ match the backness of G, as seen in [xwɑ] and [ɕja]. However, in GVX combinations, the postvocalic element takes precedence in local assimilation.

2.2 Diminutive formation in West Huozhou

Diminutives in Huozhou Chinese can be formed either through subtraction or reduplication. Subtraction in Huozhou is usually referred to as "rime change" in the literature, while reduplication is a common method for forming diminutives in Chinese dialects. Speakers also express smallness or cuteness in alternative ways, such as by adding the morpheme *xiao* [ɕjɑw^{33}] 'small' directly before a noun. The data presented in this paper are primarily from the author's consultation with three native speakers (one male and two females, aged between 50 and 65) from the town of Bailong, a representative area of West Huozhou. Additional information can be found in the Appendix. This paper focuses on the subtracted forms.

The pattern of subtraction can be generalized as segment deletion with feature movement. Only syllables that end with [j], [w], or [ŋ] can undergo subtraction, specifically those with the combinations listed in (3).[2] These observations align with descriptions in previous literature, despite some notational and sub-dialectal variations, as well as some minor divergences, which will be discussed shortly.

The examples of diminutive subtraction are given in (5)–(7), and additional examples are provided in the Appendix. As mentioned earlier, the subtracted forms are represented with a length mark to indicate bimoraicity, since they still carry a full tone. This is a phonological assumption based on previous studies of Chinese languages (§2.1), and a comprehensive phonetic study is beyond the scope of this paper. In addition, most Chinese morphemes are monosyllabic, and some morphemes can stand alone as words (see Packard 2015 for relevant discussions on free and bound roots in Chinese). Thus, in the following examples, the subtracted form of certain nouns can be used independently (e.g., 5b), while for others, the subtracted form is more commonly seen in compounds (often disyllabic) that end with those nouns (e.g., 5a). The gloss of each compound is provided in a separate column when needed, along with the literal gloss in parentheses. Note that underlying /a/ surfaces as [ɑ] when followed by [ŋ] in a closed syllable or preceded by [w] in an open syllable, while it surfaces as [a] elsewhere; some tonal changes will also be discussed shortly.

(5) Diminutive subtraction in Huozhou: *ŋ-ending* nouns

	Noun	Gloss	Subtracted	Gloss of Compound
a.	[kɑŋ^{11}]	'pole'	[lɑŋ^{35}.kaː^{55}]	'railing' (railing-pole)
b.	[p^{h}jɑŋ^{53}]	'braid'	[p^{h}jaː^{53}]	
c.	[kwɑŋ^{33}]	'house'	[fɑŋ^{53}.kwɑː^{53}]	'restaurant' (meal-house)
d.	[tɕ^{h}ɥɑŋ^{11}]	'circle'	[ɥɑŋ^{11}.tɕ^{h}ɥaː^{33}]	'circle' (round-circle)
e.	[ɥɑŋ^{53}]	'courtyard'	[ɥaː^{53}]	
f.	[ʐəŋ^{35}]	'kernel (nut)'	[ʐuː^{35}]	
g.	[kuŋ^{55}]	'stick'	[kwaj^{55}.kuː^{33}]	'crutch' (crutch-stick)
h.	[tɕ^{h}yŋ^{35}]	'skirt'	[wej^{35}.tɕ^{h}yː^{55}]	'apron' (enclose-skirt)

[2] Similar to Standard Chinese, morphemes in Huozhou are typically monosyllabic.

(6) Diminutive formation in Huozhou: *j-ending* nouns

	Noun	Gloss	Subtracted	Gloss of Compound
a.	[saj^{11}]	'sieve'	[saː^{53}]	
b.	[taj^{55}]	'sack'	[pu^{55}.t^{h}aː^{33}]	'cloth sack' (cloth-sack)
c.	[pej^{53}]	'generation'	[i^{35}.puː^{53}]	'the same generation' (one-generation)
d.	[k^{h}wej^{11}]	'container'	[k^{h}uː^{11}]	
e.	[tʂ^{h}wej^{35}]	'hammer'	[tʂ^{h}uː^{35}]	

(7) Diminutive formation in Huozhou: *w-ending* nouns

	Noun	Gloss	Subtracted	Gloss of Compound
a.	[ŋɑw^{53}]	'quilted coat'	[ŋoː^{53}]	
b.	[tsɑw^{33}]	'grass'	[tsoː^{53}]	
c.	[tɕ^{h}jɑw^{35}]	'stick'	[tɕ^{h}ɥoː^{35}]	
d.	[tɕ^{h}jɑw^{53}]	'sedan'	[tɕ^{h}ɥoː^{53}]	
e.	[t^{h}ow^{53}]	'bean'	[t^{h}uː^{53}]	
f.	[t^{h}ow^{35}]	'head'	[tʂəŋ^{55}.t^{h}uː^{33}]	'pillow' (rest-head)
g.	[low^{33}]	'basket'	[luː^{53}]	

In the examples above, the diminutive of a noun can be formed by deleting the last non-vocalic segment (post-nuclear glide or nasal), accompanied by featural changes of the nuclear vowel in some cases. This pattern is evident for nouns ending in [ŋ] and [j] in (5) and (6).[3] When a noun ends with a labiovelar glide [w], the subtracted forms remain regular, involving segment deletion along with some featural changes, as shown in (7).

Regarding the featural changes, one highlighted pattern is mid-vowel raising; specifically, mid-vowels are always raised to [u] or [y], as in (5f), (6c–e), and (7e–g). Additionally, although high vowels retain their height after subtraction, the low vowel [ɑ] (/a/) can be raised to the mid [o] when it is followed by [w] in the same syllable, as in (7a–d). This differs from the examples in (5) and (6), where the low vowel remains low in the diminutive form (e.g., [saj^{11}] → [saː^{11}], [p^{h}ɑŋ^{35}] → [p^{h}aː^{35}]).

Most subtracted forms still carry the same lexical tone as their non-diminutive counterparts and can be used independently. Therefore, the subtracted forms are viewed as bimoraic syllables on the surface, following the discussion in §2.1. However, in some cases, the tone of the subtracted form becomes a falling tone, as in (6a), (7b), and (7g). This tonal change was also reported in

[3] There is a difference between the author's findings and the documentation in Feng & Zhao (2014). While [-əŋ] is sometimes changed to [-ɯ] in Feng & Zhao (2014: 29), the speakers in this study produced [-u] instead.

previous literature (e.g., Feng & Zhao 2014), but it does not seem to be systematic and requires further investigation. Further, the tones in compounds are subject to tone sandhi rules, as seen in (5a), (5d), and others.

As mentioned earlier, another way to form a diminutive in Huozhou Chinese is through reduplication, and in most cases, there are no significant semantic differences between the reduplicated forms and those formed through subtraction. Besides Huozhou Chinese, reduplication is a common way to express smallness or cuteness in many Chinese dialects, particularly those spoken in Shanxi Province (Shen 2003). Even verbal reduplication in Standard Chinese and many dialects serves a diminutive function, expressing actions of short duration or politeness (Li & Thompson 1981; Audring et al. 2021). Although the issue of reduplication will not be the focus of this paper, some examples are given in (8) for illustrative purposes.[4] The reduplicant is always underlined. Relevant discussion can be found in §3.4.

(8) Diminutive reduplication in Huozhou

	Noun	Gloss	Reduplicated	
a.	[pʰjɑŋ^{53}]	'braid'	[pʰjɑŋ^{53}.pʰja^{11}]	ŋ-ending
b.	[ʐəŋ^{35}]	'kernel (nut)'	[ʐəŋ^{35}.ʐu^{55}]	
c.	[tɕʰyŋ^{35}]	'skirt'	[tɕʰyŋ^{35}.tɕʰy^{55}]	
d.	[taj^{55}]	'sack'	[taj^{55}.ta^{33}]	j-ending
e.	[pej^{11}]	'cup, mug'	[pej^{11}.pu^{33}]	
f.	[kʰwej^{11}]	'container'	[kʰwej^{11}.kʰu^{33}]	
g.	[tɕʰjɑw^{35}]	'stick'	[tɕʰɥoː^{35}.tɕʰɥo^{55}]	w-ending
h.	[low^{33}]	'basket'	[luː^{33}.lu^{33}]	
i.	[ljow^{35}]	'glass ball'	[lyː^{35}.ly^{55}]	

Diminutive reduplication in Huozhou is partial reduplication (8a–f), except for nouns ending in [w] (8g–i). Additionally, since diminutive subtraction only occurs for nouns that end in [ŋ], [j], or [w], diminutive forms for nouns without endings can be expressed through full reduplication, as shown in (9). As a side note, the symbols [ɿ] and [ʅ] are not standard IPA symbols, but they are commonly used to represent "apical vowels" in various Chinese languages and dialects (see Lee & Zee 2017 for additional information).

[4] Following previous research (e.g., Duanmu 1993; 1999; 2007), the reduplicant in nominal reduplication is treated as a monomoraic and neutral-toned weak syllable.

(9) Diminutive formation of open-syllable nouns

	Noun	Gloss	Diminutive	
			Reduplicated	Subtracted
a.	[tsɿː^{11}]	'tree branch'	[tsɿː^{11}.tsɿ^{33}]	–
b.	[tsʰaː^{11}]	'pocket'	[tsʰaː^{11}.tsʰa^{33}]	–
c.	[ŋɤː^{35}]	'moth'	[ŋɤː^{35}.ŋɤ^{55}]	–
d.	[tɕjaː^{11}]	'home'	[tɕjaː^{11}.tɕja^{33}]	–
e.	[xwɑː^{11}]	'flower'	[xwɑː^{11}.xwɑ^{33}]	–

Summing up, Huozhou Chinese uses subtraction to form diminutives, which is unique among Chinese dialects. Subtraction in Huozhou Chinese is a regular process where closed syllables are changed into open ones, as summarized in (10). The patterns in parentheses have been reported in other works but are less common in the author's fieldwork. The alternation between [a] and [ɑ] is due to local assimilation.

(10) Patterns of diminutive subtraction

Ending	**Low Vowel**		**Mid/High Vowel**	
[-ŋ]	[ɑŋ] → [aː] [iɑŋ] → [jaː]	[wɑŋ] → [wɑː] [yɑŋ] → [ɥaː]	[əŋ] → [uː] [iŋ] → [iː]	[uŋ] → [uː] [yŋ] → [yː]
[-j]	[aj] → [aː]	[waj] → [wɑː]	[ej] → [uː]	[wej] → [uː]
[-w]	[ɑw] → [oː]	[jɑw] → [ɥoː]	[ow] → [uː]	([jow] → [yː])

Overall, the patterns of diminutive formation can be categorized into *changes at the prosodic level* and *changes at the segmental level*. At the prosodic level, if a noun ends with [ŋ], [j], or [w] (except for [-jaj]),[5] the last non-vocalic element can be deleted to form a diminutive, as in (5), (6), and (7). At the segmental level, regular phonological processes accompany deletion, and two notable characteristics are observed. First, stepwise raising occurs in the diminutive form when the nuclear vowel is followed by [w] in a noun. The vowel is raised from low to mid ([ɑ] → [o]) or from mid to high ([o] →[u] or [y]). Second, although vowel quality alternations also occur in j-ending and ŋ-ending nouns, only mid vowels are raised in the diminutive form (e.g., [ə] → [u]) while the low vowel remains low. Both segment deletion and featural changes warrant analyses, and the underlying representations and grammar that lead to this unique phenomenon are the main puzzles. These will be addressed in the following sections.

[5] The syllables with [-jaj] are very rare among Chinese dialects in general.

2.3 Additional issues with the patterns

Before concluding the data description, several additional issues are addressed to provide a thorough understanding of the dataset.

The patterns described above generally align with the descriptions of Huozhou Chinese in previous literature. However, it is noteworthy that some discrepancies exist between the data presented in this paper and those described in earlier works, likely due to idiosyncratic variations. In previous studies, the featural changes of [ow] (/əu/) and [jow] (/iəu/) after subtraction have been described in various ways, such as [ɯ] and [iɯ] (Feng & Zhao 2014), [ɯ]/[ʉ] and [y] (X.Tian 1992), and [u] and [y] (J.Tian 2009). The data presented in this paper are consistent with the descriptions in J. Tian (2009). Nevertheless, despite the differences in featural changes, the general patterns of subtraction remain consistent across these descriptions. Moreover, although this paper mainly discusses the variety of West Huozhou, diminutives in other subvarieties of Huozhou Chinese can also be formed by dropping the final segment of a syllable (Feng & Zhao 2014).

Many aspects of Huozhou diminutive subtraction are consistent with the findings of previous studies on subtractive morphology. As Dressler (2000) and Manova (2011; 2020) point out, the deleted element in subtractive morphology is often the final segment of a morphological base, which aligns with the observations in Huozhou.[6] In addition, diminutive formation is a common function of subtraction, as has also been observed in other language families, such as the Slavic languages (Manova 2011). Regarding the typology of subtraction, Manova (2011; 2020) concludes that root-based subtraction is rare, and this conclusion also applies to the patterns reported in this paper. Some cases of subtraction in Huozhou appear to act on monosyllabic noun roots, such as (5b), (6a), and (7a), but these are always free roots; namely, those that can stand alone as words. Therefore, the data of Huozhou do not contradict the insights of previous studies.

Regarding the relation between language types and subtraction, subtraction in Huozhou, an analytic language, supports Manova's (2011: 173–174) insight that such a relation is questionable, despite Dressler's (2000: 585) speculation that subtraction does not occur in non-agglutinating or non-fusional languages. Although subtraction has been observed in various language families, including Romance, Germanic, and Slavic languages (Manova 2020: 12, and references therein), Chinese dialectologists have well-documented the subtractive patterns of Huozhou Chinese. However, these patterns have not received much attention from theorists,

[6] Some researchers (e.g., Fábregas & Scalise, 2012: 62–63) view patterns involving the deletion of a single segment as evidence for a-morphous morphology and reject the item-and-arrangement approach. Specifically, in cases of segment deletion, such as Huozhou, the deleted material does not carry any meaning; in other words, the deleted portion is not identified with an existing morpheme. Nevertheless, Manova (2020: 4–8) argues that the deletion of a single segment does not necessarily make morphology a-morphous, although the subtracted material often coincides with a morpheme (e.g., Manova 2011; 2020). I thank a reviewer for pointing out this issue.

likely because the descriptive works are written in Chinese, and the relevant discussions are not framed in contemporary linguistic terminology.

Finally, subtractive morphology should always be distinguished from templatic truncation. For instance, Bat-El (2002) distinguishes between “true truncation” and “fake truncation,” with subtractive morphology regarded as “true truncation.” In Bat-El (2002), “true truncation” is morphologically motivated and a-templatic. In other words, the remaining portion is not necessarily limited to a designated phonological unit, such as a foot or a syllable. In contrast, “fake truncation” results from the imposition of a prosodic template. In the case of Huozhou diminutive formation, the observed pattern should be treated as subtractive, or “true truncation”, since segment deletion clearly changes the meaning of the nouns. It is morphologically driven and occurs only when expressing diminutives.

3 Diminutive subtraction as floating mora affixation

In this section, I will demonstrate how Huozhou diminutive subtraction can be attributed to floating mora affixation with featural movement, thereby supporting the item-and-arrangement approach. In what follows, the proposal will be laid out in §3.1. The analysis in §3.2 focuses on subtraction, while §3.3 deals with featural changes. Although this paper focuses on the subtractive patterns, the reduplicative patterns will be briefly discussed in §3.4. Finally, a summary is given in §3.5.

3.1 Proposal

The item-and-arrangement approach builds on the standard view of morphemes inherited from structuralism, and a morpheme is defined as the smallest unit in a language that relates meaning and form (Anderson 1992: 48–51).[7] The key proposal of the current analysis is that the underlying phonological form of the diminutive morpheme is a floating mora that has not been integrated into a prosodic structure, and subtraction is a way to repair the floating mora. Subtractive morphology is rare in general and challenges the item-and-arrangement approach where addition is the only way of morphological operation. Stonham (1994) argues that the cases of so-called subtractive morphology are all subject to reanalysis in support of the proposal of Combinatorial Morphology. For instance, the masculine form of French adjectives is not necessarily derived from the corresponding feminine form via segment deletion (e.g., [sɔt]

[7] There are various alternative ways of defining *morpheme*. For example, Aronoff & Fudeman (2011: 266) define *morpheme* as “a word or a meaningful piece of a word that cannot be divided into smaller meaningful parts.” Similarly, Haspelmath & Sims (2010: 335) define *morpheme* as “the smallest meaningful part of a linguistic expression that can be identified by segmentation; a frequently occurring subtype of morphological pattern.” These definitions are more theory-neutral and are also compatible with the item-and-process approach. I thank a reviewer for pointing out the importance of properly defining *morpheme*.

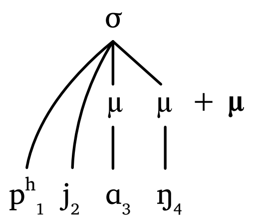
'drunk.FEM' ~ [so] 'drunk.MASC'). In an alternative analysis, the underlying form of the base consists of a coda but it is only protected in the feminine form due to the addition of a schwa-like vowel. In contrast, the consonant is dropped to favor a coda-less syllable structure in the masculine form. To some extent, the proposal of floating mora affixation for Huozhou diminutive formation also aligns with this idea. Similar to the reanalysis of French adjectives, the diminutive form in Huozhou can also be viewed as a result of affixation involving a phonological element, which is a mora in this case. The mora affixation process would lead to segment deletion in order to meet the requirements of prosodic well-formedness.

In the framework of Optimality Theory (Prince & Smolensky 1993/2004), the idea that subtractive morphology can be reanalyzed as affixation is in line with the proposals of Trommer (2011), Bye & Svenonius (2012), Trommer & Zimmermann (2014), Zimmermann (2017), and Köhnlein (2018), among others. Although previous analyses involve different theoretical apparatuses (e.g., Colored Containment versus Correspondence Theory), the core insight is consistent; namely, subtractive morphology can emerge as a way of realizing affixal features or prosodic nodes to satisfy phonological well-formedness.[8] Additionally, the treatment of various morphophonological patterns in Chinese dialects as the affixation of a floating mora and/or floating features has also been explored in Lin's (1989; 1993; 2001; 2004; 2010; 2022) work.

With the underlying affixal mora being proposed, the representations of Huozhou diminutive formations are in (11). The subscript digits indicate input-output segment correspondence. Note that the alternation between [a] and [ɑ] is attributed to local assimilation in backness.

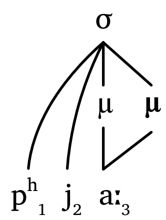
(11) Diminutive subtraction as floating mora affixation (tones are not marked)

a. Underlying: mora affix



/p^{h}jɑ$^{\mu}$ŋ$^{\mu}$ + **μ**/ 'braid + DIM'

b. Surface: subtraction



[p^{h}jaː$^{\mu\boldsymbol{\mu}}$] 'braid.DIM'

[8] This proposal follows the insights of Bye & Svenonius (2012) and Trommer & Zimmermann (2014), which suggest that subtractive patterns are epiphenomenal and can be treated as additive. These insights inherit the basics of Distributed Morphology (Halle & Marantz 1993), an item-and-arrangement model. Distributed Morphology assumes that word formation results from syntactic operations and, therefore, morphology is limited to concatenation (see Siddiqi 2018 for an overview). A key difference between Distributed Morphology and the current proposal is that, in Distributed Morphology, "morphemes" contain no phonological forms and are only abstract formal features (feature bundles), with phonological forms inserted into the structure through Vocabulary Insertion (Siddiqi 2018: 147–151). The current analysis focuses on the phonological grammar at PF, implemented within Optimality Theory. The architecture of the syntax-phonology interface and the connection between Distributed Morphology and OT are discussed in Bye & Svenonius (2012).

In (11), as well as in the analyses below, the phonological form of the diminutive morpheme is indicated by a mora in bold for the sake of readability. When the affixal mora is forced to be incorporated into the non-diminutive syllable, as in (11a), it triggers segment deletion (/ŋ$^{\mu}$/) and vowel lengthening, as in (11b), in order to maximally preserve the floating mora and maintain a well-formed bimoraic syllable at the same time.

However, one concern regarding a synchronic analysis of diminutive subtraction pertains to its nonproductivity. Both Dressler (2000: 585) and Manova (2020: 14) mention that subtraction is cross-linguistically rare and nonproductive. Diminutive subtraction in Huozhou also seems nonproductive based on the number of subtracted forms reported in the Appendix. Nevertheless, nonproductive patterns do not necessarily indicate that subtraction is absent from synchronic grammar. In an artificial language learning experiment (Kapatsinski 2017), evidence showed that some participants can generalize and learn a subtraction rule, suggesting that the subtractive forms are computable rather than merely stored (see also Kurisu 2020). Furthermore, Jackendoff & Audring (2018: 399–400) argue that learners can learn nonproductive patterns. Learners cannot initially determine whether a given pattern is fully productive, and therefore, they may form hypotheses of rules based on all their observations. Even when the pattern is later determined to be nonproductive during the learning process, the hypotheses will not be completely removed from the brain (see also Köhnlein 2018: 623–624 for a discussion of subtractive patterns in the plural formation of German dialects). Therefore, although Huozhou diminutive subtraction may not be fully productive, it still warrants a synchronic analysis.

Before moving to the analyses, the illustration in (11) suggests that the prosodic structure of the root is established at the stage of mora affixation. Various strategies have been adopted to ensure that prosodification occurs at an earlier stage. In the current analysis, I use Stratal OT (Kiparsky 2000; 2003; 2010; 2015; Trommer 2011; Bermúdez-Otero 2018, among others) to model this layering effect. The core principle of Stratal OT is that word construction operates across multiple levels, which are organized in a feed-forward manner. Each level has its own phonological grammar, structured by a parallel ranking of constraints. Specifically, I assume the noun root has already undergone a phonological cycle for optimization and prosodification.[9] To maintain focus and conciseness, I will not show the process of moraification in the following analyses. Instead, the input in the subsequent tableaux will contain a root with an established moraic structure, plus an affixal mora representing the diminutive morpheme, as in the structure of (11a).

3.2 Subtraction as floating mora affixation

As far as I am aware, the analysis of subtraction as mora affixation has not been explored in any non-Optimality-Theoretic ways. However, Lin's (1993) pioneering work on Chinese

[9] See Bermúdez-Otero (2012: 57) for a similar treatment.

morphophonology discusses a comparable case from Huojia Chinese, where segmental merger is driven by mora affixation. Lin (1993: 670–672) provides a rule-based analysis using the Affix Manifestation Principle (AMP), which states that "the effect of affixation of a phonologically expressed morpheme is always manifested" (Lin 1993: 656). Specifically, in this dialect, an affix is treated as a mora that can further be realized as a schwa ([$ə^{μ}$]). Following the requirements of the AMP and templatic constraints, the affixal mora (schwa) is incorporated into the stem and re-prosodified, resulting in the application of a segmental merger rule $V_1X_2]_σ → V_{1,2}]_σ$. An example of this derivation is as follows: /u_1n_2 + μ/ → [u_1n_2 + $ə_3$] → [$w_1\tilde{ɛ}_{2,3}$], where the segmental merger rule in the last step is driven by the templatic requirement of Chinese syllables (see also Lin 2022: 226–228; further discussion is in §4 of this paper).

Although mora affixation leads to segmental merger rather than subtraction in Lin's (1993) rule-based analysis, the core insights can also be applied to the case of Huozhou diminutive subtraction. In this paper, I analyze the subtractive pattern in Huozhou as segment deletion, driven by the prosodification of the affixal mora, which is implemented in OT. Next, I will first address the analysis of segment deletion, leaving the discussion of featural changes for §3.3.

A floating prosodic unit in the input is considered marked, and the general driving force to repair such a marked structure is the constraint HEADEDNESS(μ), which is based on HEADEDNESS(X) (following Selkirk 1995).[10] This constraint requires that every mora in the output be associated with segments. Additionally, MAXFLOAT enforces the presence of the affixal floating mora in the output (Wolf 2007). These constraints are defined in (12).

(12) a. **MAXFLOAT**: Assign a violation mark for each floating autosegment in the input that does not have a correspondent in the output.

b. **HEADEDNESS(μ)**: Assign a violation mark for every mora in the input that does not dominate a segment.

For an input such as /$p^hjɑ^μŋ^μ$ + **μ**/, a straightforward way to satisfy HEADEDNESS(μ) is to dock the floating mora onto an existing segment, which would cause either vowel lengthening or gemination. Both scenarios will violate $*σ_{μμμ}$, which penalizes a super-heavy trimoraic syllable, as defined in (13a). If the floating mora is deleted to satisfy HEADEDNESS(μ), a violation of MAXFLOAT will be incurred. Note that MAXFLOAT is not equivalent to MAX-μ (13b). One violation of MAXFLOAT entails one violation of MAX-μ, but not vice versa. These constraints can drive coda deletion at the expense of MAX-S (13c). A tableau is provided in (14). Tones are not marked in the analysis.

[10] Though Selkirk's (1995) HEADEDNESS(X) did not include mora, the constraint HEADEDNESS(μ) is used after McCarthy et al (2012: 198–199).

(13) a. ***σ**$_{\mu\mu\mu}$: Assign a violation mark for any trimoraic syllable.
b. **Max-μ**: Assign a violation mark for every input mora that has no output correspondent.
c. **Max-S**: Assign a violation mark for every input segment that has no output correspondent.

(14) Input /p^{h}jɑ$^{\mu}$ŋ$^{\mu}$ + **μ**/, 'braid + dim'

		p^{h}jɑ$^{\mu}$ŋ$^{\mu}$ + **μ**	MaxFlt	Hd(μ)	*σ$_{\mu\mu\mu}$	Max-S	Max-μ
☞	a.	p^{h}jaː$^{\mu\mu}$				1	1
	b.	p^{h}jɑː$^{\mu\mu}$ŋ$^{\mu}$			1W	L	L
	c.	p^{h}jɑ$^{\mu}$ŋ$^{\mu}$ + **μ**		1W		L	L
	d.	p^{h}jɑ$^{\mu}$ŋ$^{\mu}$	1W			L	1

The winner (14a) realizes the floating mora by docking it onto the nucleus vowel, and the deletion of the coda in (14a) avoids a super-heavy syllable at the expense of Max-S and Max-μ, compared to (14b). The floating mora in (14c) does not dominate any segments, thus violating Headedness(μ). Finally, the candidate in (14d) fails to preserve the floating mora and is ruled out by MaxFloat. In summary, MaxFloat and Headedness(μ) drive the realization of the floating mora, while *σ$_{\mu\mu\mu}$ motivates coda deletion in the output. Note that the unfaithful mapping between /ɑ/ and [a] in (14a) arises from the markedness restriction that a low back vowel cooccurs only with a velar nasal or a labiovelar glide.

However, a challenge is posed by two additional candidates, [p^{h}jɑ$^{\mu}$ŋ$^{\mu}$] and [p^{h}jɑ$^{\mu}$ŋ$^{\mu}$], where the affixal mora replaces an existing mora in the input. These two candidates harmonically bound the expected winner [p^{h}jaː$^{\mu\mu}$], since neither candidate violates Max-S. I propose that the candidates with mora replacement are ruled out by NoVacuousDocking(μ). This constraint is an extension of NoVacuousDocking(F) proposed by Wolf (2007), which requires that "floating features cannot dock onto segments that already bore the same feature value in the input" (Wolf 2007: 21). In terms of moras, when a floating mora replaces an existing mora in the output, it is considered vacuous. In other words, the docking of the floating mora is expected to manifest in the output structure and cause phonological consequences, which is comparable to Lin's (1993) Affix Manifestation Principle. The proposed NoVacuousDocking(μ) is defined as follows:[11]

(15) NoVacuousDocking(μ)
Assign a violation if an existing mora is replaced with a floating mora such that the segment to which the floating mora is associated and its corresponding input segment are associated with an equal number of moras. ("Do not replace an existing mora with the floating mora.")

[11] See Saba Kirchner (2010: 48–49) for a different definition of NoVacuousDocking(μ).

With the constraint NoVacuousDocking(μ), the candidates [p^hjɑ$^{\mu}$ŋ$^{\boldsymbol{\mu}}$] and [p^hjɑ$^{\boldsymbol{\mu}}$ŋ$^{\mu}$] can be ruled out, as shown in (16).

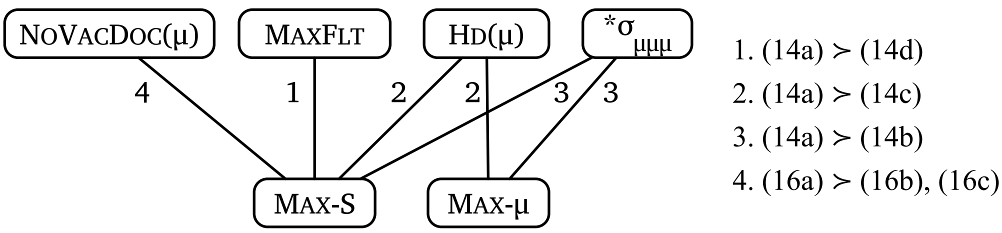
(16) Input /p^hjɑ$^{\mu}$ŋ$^{\mu}$ + **μ**/, 'braid + DIM'

	p^hjɑ$^{\mu}$ŋ$^{\mu}$ + **μ**	NoVacDoc(μ)	Max-S
☞	a. p^hjaː$^{\mu\boldsymbol{\mu}}$		1
	b. p^hjɑ$^{\mu}$ŋ$^{\boldsymbol{\mu}}$	1W	L
	c. p^hjɑ$^{\boldsymbol{\mu}}$ŋ$^{\mu}$	1W	L

The ranking established in (16) is NoVacDoc(μ) >> Max-S. The constraint NoVacDoc(μ) is violated by (16b) and (16c), where the floating mora replaces an existing mora in the input.

The constraints introduced so far can generate subtraction, and their interactions are summarized in (17). The affixal mora is forced to be realized on the surface, and the pressure of $*\sigma_{\mu\mu\mu}$ and NoVacDoc(μ) triggers segment deletion. Although there are other ways to repair a floating mora, those repair strategies can be suppressed by ranking relevant faithfulness constraints higher than Max-S and Max-μ. For example, the floating mora can also be docked on an epenthetic segment, but this will violate the higher-ranked Dep-S. The position of various faithfulness constraints determines the specific phonological instantiation of mora affixation (see §4 for additional discussion).

(17) Hasse diagram



Although the ŋ-ending noun [p^hjɑŋ^{53}] 'braid' is used for demonstration above, this same analysis applies to j- and w-ending nouns as well. It is noteworthy that featural changes, such as mid-vowel raising, are regulated by a separate set of constraints, which is the focus of the following section. The current analysis centers on how mora affixation leads to segment deletion in the diminutive form.

One reviewer asks whether it is reasonable for a mora in the root (e.g., /ŋ$^{\mu}$/) to delink and give way to a floating one. Additionally, deleting a mora in the root also incurs violations of Max-Link, which evaluates association lines (Morén 2001), thus suggesting the ranking MaxFloat >> Max-Link. Although the current analysis might seem counterintuitive, particularly given the ranking schema FAITH$_{Root}$ >> FAITH$_{Affix}$ (McCarthy & Prince 1994), it is not uncommon to protect and realize a floating mora by deleting others in the root. A relevant example can be found in the weight polarity of Diegueño. For instance, /saːw/ becomes [saw] to form the plural (Topintzi 2008: 512). Topintzi (2008: 512–513) treats the plural marker in Diegueño as a floating mora that is realized by the force of MAXFLOAT. A tableau is given in (18) to illustrate the evaluation (adapted from Topintzi 2008: 513; μ_3 is the floating mora).

(18) Input /sa$^{\mu\mu}$w + μ/, 'eat + PL'

		sa$^{\mu1\mu2}$w + μ_3	*$\sigma_{\mu\mu\mu}$	MAXFLT	P-DEP-μ	MAX-μ
☞	a.	sa$^{\mu3}$w				2
	b.	sa$^{\mu1\mu2\mu3}$w	1W			L
	c.	sa$^{\mu1\mu2}$w		1W		1L
	d.	sa$^{\mu1\mu3}$w			1W	1L

Among the losing candidates, the superheavy syllable *[sa$^{\mu1\mu2\mu3}$w] (18b) is banned by *$\sigma_{\mu\mu\mu}$. The candidate [sa$^{\mu1\mu2}$w] (18c) is ruled out by MAXFLOAT since it does not preserve the floating mora μ_3. As for *[sa$^{\mu1\mu3}$w] (18d), it is ruled out by P-DEP-μ, which penalizes any *non-positional* μ-licensed mora in the output that does not have a correspondent in the input (Topintzi 2008: 513). Specifically, a non-positional μ-licensed mora refers to a mora that dominates a segment along with other moras, and its input correspondent must also be associated with the same segment in the output (Topintzi 2006: 85). Thus, μ_3 in *[sa$^{\mu1\mu3}$w] violates this requirement because it is a non-positional μ-licensed mora and does not link to /a/ in the input. However, the winner [sa$^{\mu3}$w] is exempt from the evaluation of P-Dep-μ because it is a *positional* μ-licensed mora, i.e., the only mora that dominates the segment.[12] The core insight of the analysis is that MAXFLOAT provides extra protection to the floating elements, allowing them to take precedence over root faithfulness.

Furthermore, prioritizing and manifesting an affix in the output aligns with Samek-Lodovici's (1993) AFFIX REALIZATION constraint and Lin's (1993) Affix Manifestation Principle (AMP). A well-established example of the AMP is *-er* suffixation in Beijing Mandarin (Lin 1989: 97–98; Lin 1993: 677, cf. Duanmu 2007: 218–221), where the *-er* suffix /ɹ/ leads to coda deletion (e.g., /pan + ɹ/ → [paɹ]). Thus, the example of *-er* suffixation in Beijing also suggests that extra faithfulness

[12] The analysis of Huozhou does not consider the constraint P-Dep-μ. If P-Dep-μ is considered, it is violated by the candidates (14a) and (16a), indicating that it should be ranked lower than MAXFLT, HD(μ), *$\sigma_{\mu\mu\mu}$, and NOVACDOC(μ) in Huozhou.

protection can be afforded to the affix, even at the cost of coda deletion in the root. The idea also applies to the Huojia case discussed in §4.

3.3 Vowel raising in subtraction

This section addresses the featural changes that occur during subtraction, in order to provide a more comprehensive analysis of the diminutive patterns and resolve the related phonological issues. The most prominent process is the change in vowel height. Some examples are repeated in (19).

(19) Vowel raising in subtraction

	Noun	Gloss	Subtracted	
a.	[pɑw^{11}]	‘bag’	[poː^{11}]	w-ending
b.	[t^{h}ow^{53}]	‘bean’	[t^{h}uː^{53}]	
c.	[saj^{11}]	‘sieve’	[saː^{53}]	j-ending
d.	[pej^{53}]	‘lifetime’	[puː^{53}]	
e.	[p^{h}ɑŋ^{35}]	‘plate’	[p^{h}aː^{35}]	ŋ-ending
f.	[p^{h}əŋ^{35}]	‘basin’	[p^{h}uː^{35}]	

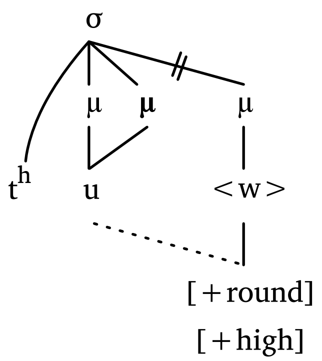
There are several remarkable patterns in (19). First, the low vowel is raised only when followed by [w] (19a), compared to [aj] (19c) and [ɑŋ] (19e). Second, the low vowel followed by [w] is raised only to mid (19a), which is an instance of stepwise raising. Third, mid vowels are raised to high regardless of the following segment, as in (19b, d, f).

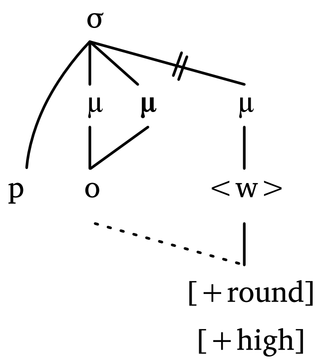
To account for these patterns, I propose that the raising of the nuclear vowel is attributed to the preservation of the [+high, +round] features of the deleted segment. The constraint MAX(feature) governs featural correspondence and preservation (McCarthy & Prince 1995; Causley 1997; Zhang 2000). Although there have been some debates over whether MAX(feature) can replace IDENT-IO(feature), some studies show the utility of both types of featural faithfulness in the grammar (e.g., Walker 2001; Wolf 2007; cf. Zhang 2000).

(20) **MAX(feature)**
Assign a violation for every input feature [F] that does not have a correspondent in the output.

The idea that vowel change is due to the preservation of features in [w] can be further illustrated in (21). When [w] is deleted, the [+round] and/or [+high] features are retained by MAX(+round) and MAX(+high), resulting in the rounding and raising of the input low vowel.

(21) Segment deletion with feature movement





a. [tʰu:$^{μ\mathbf{μ}}$] (from /tʰo^{μ}w^{μ} + **μ**/) b. [po:$^{μ\mathbf{μ}}$] (from /pɑ^{μ}w^{μ} + **μ**/)

In (21a), the main motivation for vowel raising is attributed to Max(feature). However, one potential concern is that the preservation of the [+round] feature in /w/ of /tʰo^{μ}w^{μ} + **μ**/ may seem vacuous because the main vowel already has a [+round] feature. Given feature correspondence, we may assume that the [+round] feature in /w/ (labeled as $[+\text{round}]_2$ in the tableau) is deleted instead of merging with the [+round] feature in /o/ (labeled as $[+\text{round}]_1$ in the tableau),[13] while the output is still generated correctly. See (22) for illustration.

(22) Input /to^{μ}w^{μ} + **μ**/ 'bean + DIM': raising and rounding

	$to_1{}^{μ}w_2{}^{μ}$ + **μ** $[+rd]_1$, $[+rd]_2$	MaxFlt $*σ_{μμμ}$ Hd(μ) NoVacDoc(μ)	Max (+rd)	Max (+hi)	Id (hi)	Id (rd)
☞ a.	$tu{:}_1{}^{μ\mathbf{μ}}$ $[+rd]_1$		1 $[+rd]_2$		1	
b.	$to{:}_1{}^{μ\mathbf{μ}}$ $[+rd]_1$		1 $[+rd]_2$	1W	L	
c.	$ti{:}_1{}^{μ\mathbf{μ}}$		2W $[+rd]_1$, $[+rd]_2$		1	1W
d.	$to_1{}^{μ}w_2{}^{μ}$ $[+rd]_1$, $[+rd]_2$	1W (MaxFlt)	L		L	

[13] An alternative is that two [+round] features merge in the output at the cost of Uniformity. However, this paper does not pursue the fusion (coalescence) analysis. In other words, Uniformity is always ranked higher.

In (22), the pressure of MAX(+round) and MAX(+high) forces the output to preserve the relevant features in the input, ruling out (22b) and (22c), respectively. Subtraction is motivated by the constraints introduced in §3.1, and (22d) is ruled out by MAXFLOAT.

However, the issue of stepwise raising in (21b) requires additional treatment. To block raising from low to high, I employ local conjunction (e.g., Smolensky 1993; Itô & Mester 1998; Łubowicz 2002; Smolensky 2006), which has been used in the analysis of stepwise raising (e.g., Kirchner 1996; Walker 2005; 2011). Local constraint conjunction forms complex constraints from simple ones. I follow the theoretical assumptions of Itô & Mester (1998), and the conjoined constraint relevant for Huozhou is IDENT-IO(high)&IDENT-IO(low), as shown in (23).

(23) **IDENT-IO(high)&IDENT-IO(low)**
Assign a violation mark if IDENT-IO(high) and IDENT-IO(low) are violated with respect to the same segment.

If a segment in the output violates both IDENT-IO(high) and IDENT-IO(low), the conjoined constraint ID(high)&ID(low) (a shorthand version) will be violated. This constraint penalizes the change from a low vowel to a high vowel, which skips one intermediate height level and changes from [+low, –high] to [–low, +high]. A tableau is provided in (24).

(24) Input /pɑ$^{\mu}$w$^{\mu}$ + **μ**/ ‘bag + DIM’: stepwise raising

		pɑ$_1^{\mu}$w$_2^{\mu}$ + **μ**	MAXFLT *σ$_{\mu\mu\mu}$ HD(μ) NOVACDOC(μ)	MAX (+rd)	ID(hi)& ID(lo)	MAX (+hi)	ID (hi)	ID (lo)
☞	a.	poː$_1^{\mu\mu}$				1		1
	b.	puː$_1^{\mu\mu}$			1W	L	1W	1
	c.	paː$_1^{\mu\mu}$		1W		1		L
	d.	pɑ$_1^{\mu}$w$_2^{\mu}$	1W (MAXFLT)			L		L

The winner (24a) raises the input low vowel to [o], violating MAX(+high) but satisfying MAX(+round). Since only IDENT-IO(low) is violated, the winner avoids violating the conjoined constraint. When the low vowel becomes high, as in (24b), it satisfies both MAX(+high) and MAX(+round) at the expense of IDENT-IO(high) and IDENT-IO(low), thereby triggering ID-IO(high)&ID-IO(low). Candidate (24c) fails to retain the [+round] feature in the output, fatally violating MAX(+round). Finally, (24d) violates the set of constraints driving subtraction.

The evaluation indicates that low vowel raising is epiphenomenal. MAX(+high) is expected to promote vowel raising, but the higher-ranked ID(high)&ID(low) blocks two-step raising of the low vowel to [u], which involves a simultaneous change of [high] and [low]. Further, MAX(+round) drives the low vowel to [o] rather than [ɒ], since the low back rounded vowel [ɒ] is not permitted in Standard Chinese (Duanmu 2007, cf. Zhang 2023b), and this restriction also holds in Huozhou Chinese.

This analysis accounts for the asymmetrical pattern of vowel raising. For j-ending and ŋ-ending nouns, vowel raising from low to high is also blocked by the conjoined constraint, while MAX(+round) does not motivate any epiphenomenal raising, as in (24). I assume the velar nasal is [+high, –low] (e.g., Hall 2007: 332).

(25) Input /pʰɑ^{μ}ŋ^{μ} + **μ**/ 'plate + DIM': no raising

pʰɑ$_1^{μ}$ŋ$_2^{μ}$ + **μ**	*ɑ]$_σ$	MAX (+rd)	ID(hi)& ID(lo)	MAX (+hi)	ID (hi)	ID (lo)
☞ a. pʰaː$_1^{μμ}$				1		
b. pʰoː$_1^{μμ}$				1		1W
c. pʰuː$_1^{μμ}$			1W	L	1W	1W
d. pʰɑː$_1^{μμ}$	1W			1		

With the same ranking, the grammar predicts no raising of ŋ-ending nouns (the same applies to j-ending nouns). The constraint ID-IO(high)&ID-IO(low) rules out (25c), while the winner (25a) is favored over (25b) by IDENT-IO(low). As a side note, a back low vowel is penalized in an open syllable (*ɑ]$_σ$) as a basic phonotactic rule in Huozhou Chinese unless it is preceded by a prenuclear glide [w]. Thus, the candidate in (25a) ends with [a].

The last issue concerning the featural changes is the raising of mid vowels in j-ending and ŋ-ending nouns, including [əŋ] (/əŋ/), [wej] (/uəi/), and [ej] (/əi/), all of which become [uː] after subtraction. In brief, the case of [əŋ] aligns with the analysis above, and the change from [əŋ] to [uː] can be attributed to the preservation of [+high] in [ŋ] after subtraction (recall that in (4), [ə] is specified [+back]), as illustrated in (26). The backness of the nuclear vowel is preserved by IDENT-IO(back).

(26) Input /pʰə^{μ}ŋ^{μ} + **μ**/ 'basin + DIM': mid vowel raising

pʰə$_1^{μ}$ŋ$_2^{μ}$ + **μ**	MAX (+rd)	MAX (+hi)	ID (hi)	ID (lo)	ID (bk)
☞ a. pʰuː$_1^{μμ}$			1		
b. pʰiː$_1^{μμ}$			1		1W
c. pʰəː$_1^{μμ}$		1W	L		

For [wej], I propose that [wej] → [uː] is conditioned by both MAX(+high) and local agreement. First, mid-vowel raising is motivated by MAX(+high) during subtraction; then, the onglide [w] can influence the nuclear vowel through local agreement. I posit that [kʰwuː] is phonetically realized as [kʰuː]. Finally, for [ej] (/əi/), the only syllable that has diminutive forms is [pej] (/pəi/), and the output [puː] is also possibly due to consonant-vowel interactions.

3.4 Reduplication in Huozhou diminutive formation

Although this paper focuses on subtraction in Huozhou diminutive formation, reduplication is another method for forming diminutives and deserves attention. This section briefly discusses the reduplicative patterns and addresses several essential issues. A comprehensive investigation of Huozhou diminutive reduplication, especially the patterns of w-ending nouns, will be undertaken in separate work.

It has been observed that, in most cases, there are no significant semantic differences between subtraction and reduplication. However, one speaker mentioned a pair that differs in meaning. The subtracted form [muː^{35}], derived from [məŋ^{35}] 'door', often refers to doors of houses or rooms, whereas the reduplicated form, [məŋ^{35}.mu^{55}], specifically indicates the door of animal stalls (e.g., pigpen). However, this semantic difference is not prevalent in the data collected, nor has it been reported in previous literature (e.g., X.Tian 1992; Feng & Zhao 2014). Additionally, the examples in the Appendix show that some nouns have both subtracted and reduplicated forms, while others have only one form commonly used by the consultants. Although the nature of these gaps requires further investigation, the grammar should at least be able to generate both forms.

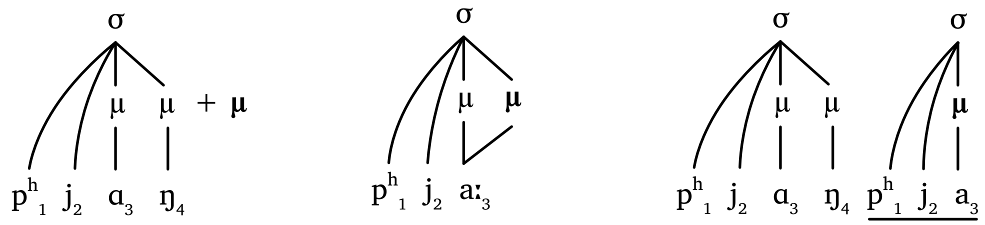
If we view subtraction and reduplication as two ways to express the same meaning, the question becomes what the underlying form would be in order to generate both forms. Reduplication is also a common method for realizing an affixal prosodic template, where the input's empty prosodic template is populated by copying the existing segments. This idea is in line with Marantz (1982) and McCarthy & Prince (1986/1996), which is developed in Saba Kirchner's (2010; 2013) Minimal Reduplication with a constraint-based implementation (cf. McCarthy et al. 2012). Given the current proposal that subtraction is attributed to mora affixation, I tentatively assume that reduplication and subtraction are surface variants of a single underlying form. As presented earlier, subtraction can be attributed to mora affixation, as in (27b); reduplication can serve as an alternative way to incorporate the floating mora into a prosodic structure; namely, by populating the empty template with existing segments, as in (27c).[14] In other words, two nonconcatenative effects can arise from the affixal mora.

---

[14] The reduplicant in nominal reduplication of Chinese dialects is often treated as a monomoraic, neutral-toned weak syllable, as noted in previous research by Duanmu (1993; 1999; 2007).

(27) Subtraction and reproduplication as floating mora affixation (tones are not marked)

a. Mora affix *b. Subtraction* *c. Reduplication*



/pʰjɑᵘŋᵘ + **μ**/ ‘braid + DIM’

[pʰjaːᵘ**ᵘ**] ‘braid.DIM’

[pʰjɑᵘŋᵘ.pʰjaː**ᵘ**] ‘braid.DIM’

Similar to the analysis of subtraction, the constraints MAXFLOAT and HEADEDNESS(μ) enforce the presence and realization of the input floating mora in the output. The faithfulness constraint that penalizes reduplication is INTEGRITY-S. Since copying segments is considered segment fission, INTEGRITY-S is violated by multiple correspondence relations between input and output segments. The constraint is defined in (28) (after McCarthy & Prince 1995: 124), and a tableau is provided in (29).

(28) **INTEGRITY-S**
Assign a violation mark for every input segment that has multiple correspondents in the output.

(29) Input /pʰjɑᵘŋᵘ + **μ**/ ‘braid + DIM’: reduplication

/pʰjɑᵘŋᵘ + **μ**/	MAXFLT	HD(μ)	$*\sigma_{\mu\mu\mu}$	MAX-S	MAX-μ	INTEG-S
☞ a. pʰjɑᵘŋᵘ.<u>pja</u>**ᵘ**						3
b. pʰjaᵘ**ᵘ**				1W	1W	L
c. pʰjɑᵘ**ᵘ**ŋᵘ			1W			L
d. pʰjɑᵘŋᵘ + **μ**		1W				L
e. pʰjɑᵘŋᵘ	1W				1W	L

The tableau demonstrates that reduplication can also be predicted using the same underlying form and the same set of constraints as in (14), along with INTEGRITY. In (29a), the floating mora is populated by copying the existing segments in /pjɑᵘŋᵘ/, violating INTEGRITY.[15] By comparing

[15] Although not shown in the tableau, the copying of the onset and non-moraic onglide can be achieved through L-ANCHOR and CONTIGUITY (Bye & Svenonius 2012). The constraint L-ANCHOR(AFFIX, STEM) requires that the left edge of the reduplicative affix is identical to that of the stem, while CONTIGUITY ensures that the copied portion forms a contiguous string. Additionally, another competing candidate $[p_1 j_2 ɑ^{μ}_3 ŋ^{μ}_4 . \underline{p_1 i_2}^{μ}]$ can be ruled out by IDENT-IO-SROLE (after McCarthy & Prince 1993; Bye & Svenonius 2012) since the segments $[j_2]$ and $[i_2]$ do not share the same syllable role.

the rankings in (29) and (14), it is evident that a reduplicated output can be generated by adjusting the ranking between Max (including both Max-μ and Max-S) and Integrity. Thus, the subtracted form (29b) will no longer be optimal if the Max constraints are ranked higher. In other words, with the same underlying representation, the surface form can vary depending on the ranking of specific faithfulness constraints.[16] In (29), with a lower ranking of Integrity, reduplication takes precedence over segment deletion.[17]

Nevertheless, a critical issue is how the grammar determines whether the output is a subtracted form (29b) or a reduplicated form (29a) if a single underlying representation is assumed. I speculate that the choice between these forms could be influenced by prosodic and/or syntactic factors.[18] Therefore, additional higher-ranked constraints might evaluate the prosodic context and favor one form over the other, but further investigation is needed to explore this hypothesis. It is not uncommon, however, for prosodic or syntactic factors to play a role in determining surface morphophonological patterns. A relevant example is *zi* affixation in the Linyi dialect of Zhongyuan Mandarin, also spoken in south Shanxi Province.[19] The phonetic form of the *zi* affix in Linyi can be either the monosyllable [təw^{0}] ("0" indicates a neutral tone) or realized as vowel lengthening, the latter of which can also be viewed as mora affixation. Based on the descriptive data in Wang (1993), the selection between these two forms is influenced by the syntactic structure, as exemplified in (30) (adapted from Wang 1993: 98). The [təw^{0}] form is used when the noun serves as the object of a VP (30a), while vowel lengthening is adopted when the noun is part of a compound (30b).

(30) *zi* affixation in Linyi dialect: root [fa^{31}] 'brush'

a.	[la^{24}	fa^{31}-təw^{0}]	b.	[xaj^{24}	faː^{31}]
	bring	brush-*zi*		shoe	brush-*zi*
	'bring a brush'			'shoe brush'	

For Huozhou diminutive formation, whether syntactic or prosodic factors determine the choice between subtraction and reduplication requires further investigation.[20] If relevant, additional markedness constraints would intervene in the current ranking and select between (29a) and (29b).

[16] The constraint Integrity is not included in (14), but it should be ranked higher than the Max constraints to block reduplication in that case.

[17] Another way to realize the floating mora is to epenthesize a default segment, which can be ruled out by the higher-ranked Dep-S constraint.

[18] An alternative solution is to assume that the Max constraints and Integrity are unordered, following Anttila's (1997) theory of Partially Ordered Constraints. However, unordered constraints are a form of cophonology, while the current proposal adheres to GNA, which explicitly opposes cophonology.

[19] *zi* is a common suffix in Chinese, but its phonological forms vary across dialects. It functions as a nominal affix and sometimes indicates diminutive, as discussed in Lin (1993: 650–651).

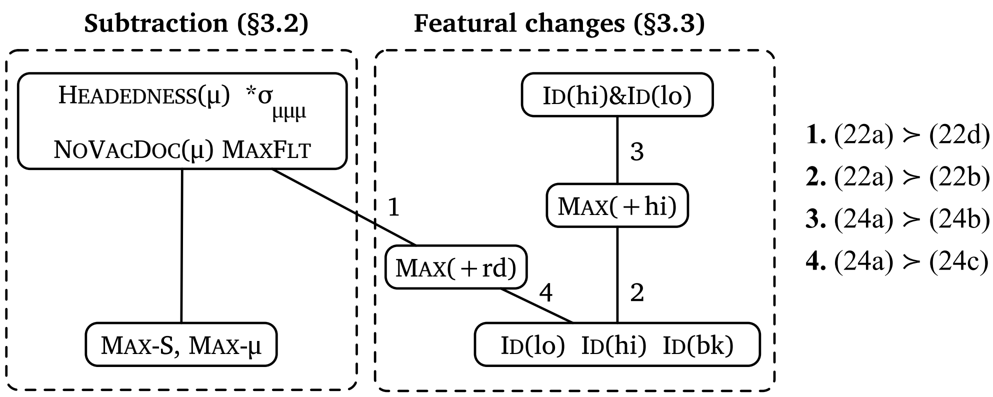
[20] Speakers may also assume two underlying representations. However, whether a single or multiple underlying forms are assumed, similar grammatical mechanisms are needed to select the surface form.

To some extent, the brief discussion on reduplication also supports an item-and-arrangement analysis for subtraction. Since both subtraction and reduplication are variations of a single morphological construction, a unified treatment of both patterns is desirable. Specifically, since various studies have successfully analyzed reduplication as the affixation of prosodic nodes (e.g., Saba Kirchner 2010; McCarthy et al. 2012), it is better to treat subtraction with the item-and-arrangement approach as well.

3.5 Summary

In sum, §3.2 demonstrates how the affixal mora is realized as subtraction under the pressure of a set of markedness constraints, including MAXFLOAT, HEADEDNESS(μ), $*\sigma_{\mu\mu\mu}$ and NOVACDOC(μ). Further, the patterns of vowel raising are discussed in §3.3 to provide a more comprehensive analysis. The constraint interactions from both sections are integrated and summarized in (31) (the phonotactic constraint $*\text{ɑ}]_\sigma$ is omitted for conciseness). Since this paper focuses on subtraction, the issues related to reduplication in §3.4 are not included in the summary below. The ranking arguments for the constraints in §3.2 are presented in (17).

(31) Hasse diagram



4 Typology of diminutive patterns in neighboring dialects

The analysis thus far demonstrates that Huozhou diminutive subtraction can be attributed to floating mora affixation. This proposal gains further support from a typological perspective. This section shows that mora affixation can result in various realizations beyond subtraction, as observed in neighboring dialects such as Heshun and Huojia. As summarized by Zimmermann (2017), research has shown that mora affixation can generate diverse surface patterns. However, if varieties of the same language family (Chinese) in a specific region exhibit both additive and subtractive patterns, it offers further support for analyzing these patterns in a unified fashion.

Diminutive subtraction in Huozhou is traditionally referred to as “rime change,” a morphophonological process where the rime part of a syllable undergoes sound changes to express meaning. It is a typical feature of many varieties of Zhongyuan Mandarin and its neighboring Jin Chinese, and most cases of rime change denote diminutives. There has been extensive descriptive work on rime change in Chinese dialects (e.g., He 1982; X. Tian 1986; Hou & Wen 1993). Lin (1989; 1993; 2001; 2004; 2010; 2022) and Duanmu (1990) have conducted systematic and in-depth formal analyses of rime change.

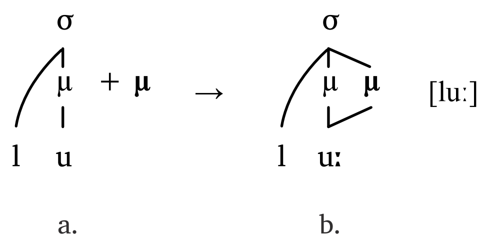
As mentioned earlier, an essential insight of Lin’s (1993, et seq.) analysis is that many rime change patterns can be treated as the affixation of a floating mora and/or floating features, which aligns with the analysis of Huozhou diminutive formation. A case study by Lin (1993) focuses on Heshun, a variety of Jin Chinese in eastern Shanxi Province, as exemplified in (32) (the superscript digits indicate tones).

(32) Heshun diminutive rime change (X. Tian 1986, cited and adapted from Lin 1993: 658)

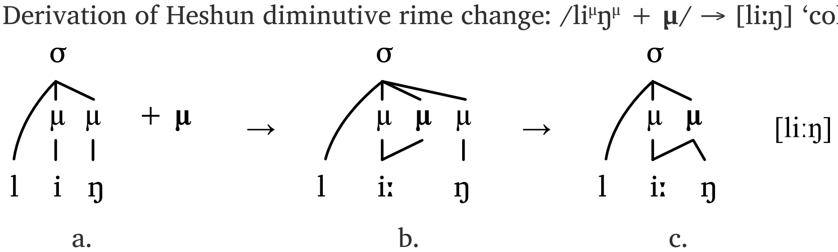
	Noun	Diminutive	Gloss		Noun	Diminutive	Gloss
a.	/lu^{22}/	[luː^{22}]	‘stove’	d.	/ɕiəu^{44}/	[ɕjəːw^{44}]	‘sleeve’
b.	/tai^{44}/	[taːj^{44}]	‘bag’	e.	/liŋ^{35}/	[liːŋ^{35}]	‘collar’
c.	/tsua^{35}/	[tswaː^{35}]	‘claw’	f.	/ʂəŋ^{31}/	[ʂəːŋ^{31}]	‘body’

The generalization from the dataset is that the nuclear vowel in each noun is lengthened to form a diminutive. Therefore, the diminutive rime change in Heshun can also be viewed as mora affixation. Lin’s (1993) analysis is presented within a rule-based framework with some constraint-based elements, and the derivations are as follows (adapted from Lin 1993: 659).

(33) Derivation of Heshun diminutive rime change: /lu$^{\mu}$ + **μ**/ → [luː] ‘stove’



(34) Derivation of Heshun diminutive rime change: /li$^{\mu}$ŋ$^{\mu}$ + **μ**/ → [liːŋ] ‘collar’



In (33a), the affixal mora is incorporated into the syllable, resulting in lengthening (33b). In (34), the incorporation of the affixal mora results in a trimoraic syllable at an intermediate stage (34b), and a template constraint $[\mu\mu]_\sigma$ motivates mora deletion and segment reassociation in (34c) (cf. Duanmu 1990: 57–58).[21] If this analysis is implemented in Optimality Theory, lengthening will be attributed to mora affixation at the expense of IDENT-IO-length. The crucial aspect of this analysis is that Heshun diminutive rime change surfaces as vowel lengthening through mora affixation.[22] Beyond Heshun, the same lengthening pattern is also found in Yuncheng (Lyu 1991), a variety of Zhongyuan Mandarin spoken in southern Shanxi Province, near Huozhou.

Another important example is the Huojia D rime change. Huojia D refers to a class of rime changes in Huojia Chinese, a variety of Jin Chinese spoken in northwestern Henan Province, which borders Shanxi Province. This type of rime change specifically applies to place names, verbs, adjectives, and adverbs (He 1982). When applied to place names (usually village names), it also conveys a diminutive and hypocoristic meaning. The patterns are provided in (35), where only *rimes* are listed. Huojia Chinese shares the same syllable structure as Standard Chinese;[23] the prenuclear glide is excluded.

(35) Huojia D rime change (adapted from He 1982: 27, see also Lin 1993: 670)

a.	*Rime*	*Rime-changed*	b.	*Rime*	*Rime-changed*	c.	*Rime*	*Rime-changed*
	[u]	[wəː]		[in]	[jɛ̃ː]		[ɑw], [ow]	[ɔː]
	[i]	[jɛː]		[un]	[wɛ̃ː]		[aj], [ej]	[ɛː]
	[y]	[ɥɛː]		[yn]	[ɥɛ̃ː]		[an]	[ɑ̃ː]
				[iŋ]	[jɔ̃ː]		[ən]	[ɛ̃ː]
				[uŋ]	[wɔ̃ː]		[ɑŋ]	[ɔ̃ː]
				[yŋ]	[ɥɔ̃ː]		[əŋ]	[ɔ̃ː]

Lin (1993) proposed that Huojia D rime change can also be treated as mora affixation, with the affixal mora realized as a default epenthetic segment [ə] (cf. Lin 2001; 2004; 2010). Therefore, for rimes containing only a monophthong, a mid vowel is inserted after rime change (35a).[24]

21 A reviewer suggests an alternative analysis: only the nuclear vowel may be moraic, and therefore, both CV and CVC syllables are monomoraic, while the rime-changed forms are bimoraic due to mora affixation. The advantage of this analysis is that there would be no intermediate stage like (34b). Admittedly, this analysis is more straightforward, but I remain conservative and follow Lin's (1993) analysis in the current discussion. Whether CV and CVC syllables are monomoraic in Heshun requires a more comprehensive investigation of its phonology.

22 One issue with the Heshun data deserves mention. Following studies of Standard Chinese, as introduced in Section 2.1, Chinese syllables that carry a full tone are assumed to be bimoraic (except for affixes), and there is no underlying length contrast in Standard Chinese. However, it appears that vowel length is distinctive in Heshun, as observed in the lengthening of vowels in the rime-changed forms (32). This property of Heshun warrants special attention.

23 Both Huojia Chinese and Standard Chinese have the maximal syllable template of CGVX, but Huojia allows a glottal coda.

24 According to He (1982), another two open rimes in Huojia, /ɿ/ and /ʅ/, become [ɐ] in D rime change. This pattern appears to be an exception so it is not discussed here. Further investigation is needed.

When a high vowel precedes a nasal coda, it is merged with the epenthetic mid vowel (35b). When a nonhigh vowel is followed by a high vowel or a nasal, segment merging occurs without epenthesis (35c).

This proposal can be realized through constraint interactions, aligning with the analysis presented in this paper. In Huojia D rime change, the affixal mora can surface as an epenthetic segment or trigger segment merger, where DEP-S and UNIFORMITY play a crucial role (compared to MAX-S and MAX-μ in Huozhou diminutive formation). A sketch of the core analysis is presented below. Several top-ranked constraints that regulate mora realization are omitted (MAXFLOAT, HEADEDNESS, *$\sigma_{\mu\mu\mu}$). The candidates in the following tableaux always preserve the floating mora, and one of the input moras is deleted to give way to the floating one (recall relevant discussions in §3.2). In both (36) and (37), the constraint *$\tilde{V}_{[+high]}$ penalizes nasalized high vowels (after Lin 2001), and UNIFORMITY bans the coalescence of input segments. The subscript digits indicate input-output correspondence.

(36) Input /u + **μ**/, /un + **μ**/

	uː$_1^{\mu\mu}$ + **μ**	NOVACDOC(μ)	*$\tilde{V}_{[+high]}$	MAX-S	DEP-S	UNIFORMITY
☞ a.	w$_1$əː$_2^{\mu\mu}$				1	
b.	uː$_1^{\mu\mu}$	1W			L	
	u$_1^{\mu}$n$_2^{\mu}$ + **μ**	NOVACDOC(μ)	*$\tilde{V}_{[+high]}$	MAX-S	DEP-S	UNIFORMITY
☞ c.	w$_1$ɛ̃ː$_{2,3}^{\mu\mu}$				1	
d.	uː$_1^{\mu\mu}$			1W	L	
e.	ũː$_{1,2}^{\mu\mu}$		1W		L	1W
f.	u$_1^{\mu}$n$_2^{\mu}$	1W			L	

(37) Input /au + **μ**/ and /an + **μ**/

	a$_1^{\mu}$u$_2^{\mu}$ + **μ**	NOVACDOC(μ)	*$\tilde{V}_{[+high]}$	MAX-S	DEP-S	UNIFORMITY
☞ a.	ɔː$_{1,2}^{\mu\mu}$					1
b.	aː$_1^{\mu\mu}$			1W		L
	a$_1^{\mu}$n$_2^{\mu}$ + **μ**	NOVACDOC(μ)	*$\tilde{V}_{[+high]}$	MAX-S	DEP-S	UNIFORMITY
☞ c.	ɛ̃ː$_{1,2}^{\mu\mu}$					1
d.	aː$_1^{\mu\mu}$			1W		L
e.	a$_1^{\mu}$ɛ̃$_{2,3}^{\mu}$	1W			1W	L

Unlike Huozhou diminutive subtraction, segment deletion is banned in Huojia, as indicated by the higher-ranked MAX-S. When the input contains a *high* vowel, as in (36), the floating mora is repaired by epenthesis (36a, c), while other candidates are ruled out by either NOVACDOC(μ) (36b, f) or *Ṽ$_{[+\text{high}]}$ (36e). In (36a), the floating mora is preserved and links to the epenthetic [ə] (marked as [əː$_2$]). Additionally, the high vowel /u$_1$/ in the input forms a glide [w$_1$] in (36a) to resolve the hiatus (*[u^μ.ə^μ]), with reassociation of the input mora (recall that moraicity does not extend to the prenuclear glide in a Chinese syllable, as introduced in §2.1). In (36b), there is no epenthesis, and the floating mora vacuously docks on the input /uː$_1$/, but it is ruled out by NOVACDOC(μ) for the same reason as in (16b, c).

The winner (36c) presents another scenario. When the input contains a nasal coda, the floating mora docks on an epenthetic vowel that merges with the nasal coda as [ɛ̃ː$_{2,3}$], ensuring the output is monosyllabic (*[u$_1^μ$.n$_2^μ$ə$_3^μ$], *[u$_1^μ$n$_2^μ$.ə$_3^μ$]). Nevertheless, (36c) does not violate UNIFORMITY because it is the epenthetic vowel that merges with the input segment. The candidate in (36d) deletes the nasal coda, violating MAX-S, while (36e) merges the input vowel with the nasal coda, violating UNIFORMITY.

In contrast, when the input contains a *nonhigh* vowel, as in (37), the surface forms show merger (coalescence) (37a, c) since an epenthetic schwa cannot be properly incorporated into the syllable. The floating mora drives coalescence (37a) instead of deletion (37b) due to the ranking MAX >> UNIFORMITY. Note that (37a) does not violate NOVACDOC(μ) because the corresponding segments are associated with different numbers of moras, according to the definition of NOVACDOC(μ) in (14). In (37e), if the floating mora docks on an epenthetic vowel that merges with the nasal coda, it violates NOVACDOC(μ). Further, nonhigh vowels are less preferred to form a glide for hiatus resolution (Casali 1996), which helps rule out forms such as *[A$_1$ɔː$_{2,3}^{μμ}$] and *[A$_1$ɛ̃ː$_{2,3}^{μμ}$] (where "A" represents a low glide; these candidates are not shown in the tableaux). This explains why epenthesis is allowed only in (35a, b), where there is an underlying high vowel. Furthermore, the output must be monosyllabic, as noted by Lin (1993), ruling out forms such as *[a$_1^μ$.w$_2$ə$_3^μ$] and *[a$_1^μ$.n$_2$ə$_3^μ$] (also not shown in the tableaux). The regulation of output shape can be achieved through constraints like ALIGN[σ] in Feng (2006).[25]

In sum, diminutive formation in Huozhou and similar processes in neighboring dialects demonstrate the versatility of mora affixation. It has been discussed in the literature that mora affixation can result in morphological lengthening, gemination, subtraction, epenthesis, and segment merger, and data from Chinese dialects in this region cover most typical realizations of mora affixation, as summarized in (38). This typology highlights a key aspect of GNA: The nonconcatenative effects of an input floating mora depend on how the phonological grammar

[25] However, this sketch does not account for an asymmetrical pattern observed in Huojia D rime change. When [ɑŋ] and [əŋ] undergo D rime change, the contrast between them is lost, and both rimes become [ɔ̃]. Lin (2010) provides an explanation, attributing it to contrast preservation.

responds (Bermúdez-Otero 2012), specifically through the relative ranking of markedness and faithfulness constraints.

(38) Realizations of mora affixation in Chinese dialects

	Realization	*Example*	*Type of faithfulness violated*
a.	subtraction	Huozhou diminutive	MAX
b.	reduplication	Huozhou diminutive	INTEGRITY
c.	lengthening	Heshun and Yuncheng diminutive	IDENT
d.	epenthesis	Huojia D rime change	DEP
e.	merger	Huojia D rime change	UNIFORMITY

The nonconcatenative patterns observed in these dialects can all be attributed to mora affixation, but the actual surface forms vary due to differences in phonological grammar. While the markedness constraints that govern the realization of the floating mora are similar, such as MAXFLOAT, HEADEDNESS(μ), NOVACDOC(μ), and $*\sigma_{\mu\mu\mu}$, the crucial distinction lies in how faithfulness constraints are ranked in each dialect. For instance, if the nonconcatenative effect manifests as subtraction, it indicates that the grammar prioritizes other faithfulness constraints over MAX.

Although the utility of prosodic node affixation (mora, syllable, foot) has been supported by various cross-linguistic phenomena (e.g., Zimmermann 2017), the typology observed in closely related Chinese dialects further shows the advantage of a unified item-and-arrangement approach to nonconcatenative morphology. With a single underlying representation, various realizations of morphological processes can be explained as the result of constraint permutation in phonology, where the ranking of faithfulness constraints plays a crucial role.

Finally, the typological evidence favors the proposal of mora affixation over an analysis with a TRUNC morpheme, as proposed by Benua (1997). Compared to the highly specialized TRUNC morpheme, the mora is a more general phonological element, and mora affixation more effectively captures the typology of diminutives in a cluster of dialects, as illustrated in this section.

5 Alternative analyses

In this section, I focus on two alternative analyses of subtractive morphology in the framework of OT, Transderivational Antifaithfulness (Alderete 1999; 2001; Horwood 2001) and Realizational Morphology Theory (Kurisu 2001), both of which are akin to item-and-process.

As put by Zimmermann (2017: 279), in both Transderivational Antifaithfulness (TAF) and Realizational Morphology Theory (RMT), subtraction and other nonconcatenative patterns result from the requirement that morphologically complex forms be phonologically distinct from their base. To some extent, both theories possess characteristics of item-and-arrangement because morphemes are regarded as entities, and nonconcatenative patterns result from affixation. However, the analyses with TAF and RMT are also considered processual, since no underlying form is assumed for the morphemes that exhibit subtraction; instead, deletion on the surface is determined solely by phonological grammar (Kurisu 2001: 265–266; Kurisu 2020: 9–11; see also Manova 2016: 17).

TAF integrates antifaithfulness constraints and Transderivational Correspondence (Benua 1997), which jointly trigger morphologically-conditioned phonological alternations. Alderete (1999, 2001) argues that every faithfulness constraint has a negative counterpart that demands deviations from the base, and antifaithfulness constraints should be considered part of Universal Grammar, alongside markedness and faithfulness constraints. Besides, antifaithfulness constraints evaluate morphological constructions with OO-correspondence. For instance, segment deletion in subtractive morphology can be attributed to ¬OO-MAX-S, as defined in (39) (following Alderete 2001: 218). To account for diminutive subtraction in Huozhou Chinese, the antifaithful constraint ¬OO-MAX-S must dominate MAX-S. An abridged tableau illustrating this alternative analysis is given in (40).

(39) **¬OO-MAX-S**
It is not the case that every segment in the input (base) has a correspondent in the output.

(40) Input /p^{h}jɑŋ + DIM/ 'braid + DIM'

	Base	p^{h}jɑŋ + DIM	¬OO-MAX-S	MAX-S
☞ a.	p^{h}jɑ$^{\mu}$ŋ$^{\mu}$	p^{h}jaː$^{\mu\mu}$		1
b.	p^{h}jɑ$^{\mu}$ŋ$^{\mu}$	p^{h}jɑ$^{\mu}$ŋ$^{\mu}$	1W	L

In (40), the constraint ¬OO-MAX-S evaluates output-output correspondence between the base and the derived output. The candidate (40b) violates the higher-ranked ¬OO-MAX-S because every segment in the input base has a correspondent in the derived output. In other words, some segments need to be deleted to satisfy the antifaithfulness constraint, making (40a) the winner.

In RMT, the driving force of nonconcatenative morphology is the constraint REALIZEMORPHEME (RM), which requires that the output be distinct from the input. The constraint is defined in (41) (Kurisu 2001: 39).

(41) REALIZEMORPHEME (RM)
Let α be a morphological form, β be a morphosyntactic category, and F(α) be the phonological form from which F(α + β) is derived to express a morphosyntactic category β. Then RM is satisfied with respect to β iff F(α + β) ≠ F(α) phonologically.

The mechanism for generating nonconcatenative morphology is to rank RM higher than specific faithfulness constraints. The ranking must be RM >> Other Faithfulness Constraints >> MAX-S for subtraction. In other words, to satisfy RM, the output must be phonologically different from the input, and the lower-ranked MAX-S indicates that deletion is the least antagonistic way to realize the input morpheme. In addition, RMT explicitly assumes indexed faithfulness constraints, and every faithfulness constraint should be indexed for each morpheme in a language (Zimmermann 2017: 284).

One major argument against these theories is that they are less restrictive than a strict item-and-arrangement model like GNA (e.g., Zimmermann 2013: 20–22; Zimmermann 2017: 309–317). In both TAF and RMT, a morpheme can, in principle, be realized through any phonological operation that makes the morphologically complex form distinct from the input. Consequently, on the one hand, the grammar needs to restrict random changes by creating an antifaithful counterpart for every faithfulness constraint, which leads to the problem of constraint proliferation.[26] On the other hand, some phonological operations involved in morpheme realization are completely unattested, leading to the issue of overgeneration.

Kurisu (2020) refutes the criticism of non-restrictiveness by pointing out that it is premature to claim overgeneration until we have a comprehensive typological study. Admittedly, it is difficult to conduct an exhaustive investigation of morphological patterns in world languages, but the case study of Huozhou favors the item-and-arrangement approach for at least two reasons. First, as discussed in §4, the types of diminutive forms in neighboring dialects are limited and restricted, falling precisely within the predictions of mora affixation. Second, many morphophonological patterns in the dialects of this region are argued to have historically involved a monosyllabic root plus a monosyllabic suffix (e.g., *-er* or *-zi*; Wang 2008: 211–216). Over time, the suffix underwent reduction and eventually merged with the root, resulting in various rime change patterns. Although the historical development of these patterns requires further research, the final stage of the reduction process may have left the original monosyllabic suffixes with no segments, leaving behind a mora template – a "degenerate affix" in Lin's (1993) terms. If this is the case, pursuing an item-and-arrangement model and analyzing these patterns as mora affixation is a reasonable approach.

Subtractive morphology can also be handled by Cophonology Theory (e.g., Orgun 1996; Inkelas & Zoll 2005). Cophonology Theory assumes the coexistence of multiple sub-grammars within a language, with specific morphological constructions subscribing to their own sub-grammars, or cophonologies. One simple way to generate subtraction is by assuming a cophonology where NOCODA >> MAX-S, specific to the DIMINUTIVE morpheme. However, the primary aim of this

[26] Similarly, in RMT, the grammar needs to restrict random changes by ranking all the other indexed faithfulness constraints above RM.

work is to offer further evidence for the item-and-arrangement approach, particularly GNA, by demonstrating that subtractive patterns can be attributed to mora affixation. A detailed comparison between GNA and Cophonology Theory is not included due to space limitations. In essence, GNA seeks to differentiate the roles of morphology and phonology, whereas cophonologies blur the boundary between these two components. A more in-depth discussion of the debate between GNA and Cophonology Theory can be found in Bermúdez-Otero (2012).

Finally, one reviewer raises a concern about the complexity of the current analysis and questions whether the alternatives might be simpler in terms of constraint interaction. In fact, both the proposed GNA analysis and the item-and-process alternatives involve complexity in different areas. As summarized in (31), the phonological grammar must account for both segment deletion and vowel raising. While the constraints and interactions for vowel raising remain unchanged, the grammar driving deletion differs due to different theoretical mechanisms. The GNA analysis introduces complexity in representation by positing a floating mora as the underlying representation of the diminutive morpheme. However, the grammar only involves regular constraints without resorting to indexation. In contrast, the item-and-process alternatives, TAF and RMT, appear simpler in representation since no underlying representation is assumed for the diminutive morpheme, as shown in (40). The issue with these theories is that they introduce additional constraints into the OT grammar and may lead to unattested predictions. Thus, it is difficult to determine whether one analysis is simpler than another. Nevertheless, the proposed GNA analysis is able to make restrictive and attested predictions, as argued in previous sections.

6 Conclusion

This paper advocates for the item-and-arrangement approach to morphology, particularly Generalized Nonlinear Affixation, using Huozhou diminutive formation as a case study. The analysis of Huozhou demonstrates the utility of mora affixation, and this proposal is further supported from a typological perspective. There are two main contributions. First, it introduces new empirical data into the discussion of subtractive morphology and gives a formal analysis of the observed patterns. While subtraction has been well-documented in language families such as Romance, Germanic, and Slavic (Manova 2020), the subtractive patterns of Huozhou Chinese have received less attention. Moreover, Huozhou's status as an analytic language supports the argument of Manova (2011) that the connection between language typology and subtraction is not straightforward.

Second, while the idea that subtraction is driven by mora affixation aligns with Trommer (2011), Trommer & Zimmermann (2014), Zimmermann (2017), and Köhnlein (2018), this paper provides cross-dialectal evidence to support this proposal. As discussed in Section 4, diminutive formation in nearby dialects exhibits various nonconcatenative effects, and an item-and-arrangement model accounts for all these possible forms. Theories related to the item-and-process approach often run into issues of non-restrictiveness, as discussed in §5. When there is no restriction on representation, the actual phonological instantiations of nonconcatenativity depend solely on grammar, and all the other possible forms need to be eliminated by indexed (anti-)faithfulness constraints. In contrast, although an item-and-arrangement model such as GNA introduces complexity in representation, it strikes a balance between representational and grammatical complexity, resulting in a more economical grammar and more restrictive predictions.

Abbreviations

DIM = diminutive, PFV = perfective, IPFV = imperfective, MASC = masculine, FEM = feminine.

Additional file

The additional file for this article can be found as follows:

- **Appendix.** Examples of Huozhou diminutive formation. DOI: https://doi.org/10.16995/glossa.15130.s1

Ethics and consent

The Institutional Review Board (IRB) designee of the University of Southern California reviewed the information pertaining to the study and concluded that the project does not qualify as Human Subjects Research.

Acknowledgements

I am grateful to the associate editor of *Glossa*, Björn Köhnlein, two anonymous reviewers, Zhang Jisheng, Zhang Hongming, Wang Xiaopei, and the participants of the 29th Annual Conference of the International Association of Chinese Linguistics for their insightful questions, comments, and suggestions. Special thanks go to Rachel Walker for her continuous support and valuable feedback throughout this research. I would also like to thank Maxwell Kaplan for his assistance with the manuscript, as well as the consultants of Huozhou Chinese for their collaboration and help. All remaining errors are my own.

Competing interests

The author has no competing interests to declare.

References

Akinlabi, Akinbiyi. 1996. Featural affixation. *Journal of Linguistics* 32(2). 239–289. DOI: https://doi.org/10.1017/S0022226700015899

Alderete, John. 1999. *Morphologically governed accent in Optimality Theory*. Amherst, MA: University of Massachusetts dissertation. DOI: https://doi.org/10.1017/S0022226700015899

Alderete, John. 2001. Dominance effects as transderivational anti-faithfulness. *Phonology* 18(2). 201–253. DOI: https://doi.org/10.1017/S0952675701004067

Anderson, Stephen. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511586262

Anttila, Arto. 1997. Deriving variation from grammar. In Hinskens, Frans & van Hout, Roeland & Wetzels, W. Leo (eds.), *Variation, change, and phonological theory*, 35–68. Amsterdam: John Benjamins. DOI: https://doi.org/10.1075/cilt.146.04ant

Aronoff, Mark & Fudeman, Kirsten Anne. 2011. *What is morphology?* (2nd ed). Malden, MA: Wiley-Blackwell.

Audring, Jenny & Leufkens, Sterre & van Lier, Eva. 2021. Small events. Verbal diminutives in the languages of the world. *Linguistic Typology at the Crossroads* 1(1). 223–256. DOI: https://doi.org/10.6092/issn.2785-0943/13427

Bao, Zhiming. 1990. *On the nature of tone*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Bao, Zhiming. 1996. The syllable in Chinese. *Journal of Chinese Linguistics* 24(2). 312–353.

Bat-El, Outi. 2002. True truncation in colloquial Hebrew imperatives. *Language* 78(4). 651–683. DOI: https://doi.org/10.1353/lan.2003.0003

Benua, Laura. 1997. *Transderivational identity: Phonological relations between words.* Amherst, MA: University of Massachusetts dissertation. DOI: https://doi.org/10.7282/T34F1PK3

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Trommer, Jochen (ed.), *The morphology and phonology of exponence*, 8–83. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199573721.003.0002

Bermúdez-Otero, Ricardo. 2018. Stratal Phonology. In Hannahs, Stephen J. & Bosch, Anna R. K. (eds.), *The Routledge handbook of phonological theory*, 100–134. Abingdon: Routledge. DOI: https://doi.org/10.4324/9781315675428-5

Bye, Patrick & Svenonius, Peter. 2012. Non-concatenative morphology as epiphenomenon. In Trommer, Jochen (ed.), *The morphology and phonology of exponence*, 427–495. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199573721.003.0013

Casali, Roderic F. 1996. *Resolving hiatus*. Los Angeles, CA: University of California dissertation. DOI: https://doi.org/doi:10.7282/T3MP5260

Causley, Trisha. 1997. Identity and featural correspondence: the Athapaskan case. In Kusumoto, Kiyomi (ed.), *Proceedings of the North East Linguistic Society* 26. 93–105. Amherst, MA: GLSA Publications.

Chao, Yuen-Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Cheng, Robert L. 1966. Mandarin Phonological Structure. *Journal of Linguistics* 2(2). 135–158. DOI: https://doi.org/10.1017/S0022226700001444

Davis, Stuart & Ueda, Isao. 2002. Mora augmentation processes in Japanese. *Journal of Japanese Linguistics* 18. 1–23. DOI: https://doi.org/10.1515/jjl-2002-0104

Davis, Stuart & Ueda, Isao. 2006. Prosodic vs. morphological mora augmentation. *Lexicon Forum* 2. 121–143.

Dressler, Wolfgang. 2000. Subtraction. In Booij, Geert & Lehmann, Christian & Mugdan, Joachim & Kesselheim, Wolfgang & Skopeteas, Stavros (eds.), *Morphology: An international handbook on inflection and word formation (Vol. 1)*, 581–587. Berlin: De Gruyter Mouton. DOI: https://doi.org/10.1515/9783110111286.1.8.581

Duanmu, San. 1990. *A formal study of syllable, tone, stress and domain in Chinese languages*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Duanmu, San. 1993. Rime length, stress, and association domains. *Journal of East Asian Linguistics* 2(1). 1–44. DOI: https://doi.org/10.1007/BF01440582

Duanmu, San. 1994. Syllabic weight and syllabic duration: A correlation between phonology and phonetics. *Phonology* 11(01). 1–24. DOI: https://doi.org/10.1017/S0952675700001822

Duanmu, San. 1999. Metrical structure and tone: evidence from Mandarin and Shanghai. *Journal of East Asian Linguistics* 8(1). 1–38. DOI: https://doi.org/10.1023/A:1008353028173

Duanmu, San. 2007. *The Phonology of Standard Chinese* (2nd ed). New York: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780199215782.001.0001

Fábregas, Antonio & Scalise, Sergio. 2012. *Morphology: From data to theories*. Edinburgh University Press. DOI: https://doi.org/10.1515/9780748656264

Feng, Guanjun. 2006. *Morpheme recognition in prosodic morphology*. Los Angeles, CA: University of Southern California dissertation.

Feng, Liangzhen & Zhao, Xueling. 2014. *Huozhou Fangyan Yanjiu* [The Study of Huozhou Dialect]. Taiyuan: Beiyue Literature & Art Publishing House.

Goldsmith, John. 1976. *Autosegmental Phonology*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

Hall, Tracy Alan. 2007. Segmental features. In de Lacy, Paul (ed.), *The Cambridge Handbook of Phonology*, 311–334. Cambridge, New York: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511486371.014

Halle, Morris & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. In Hale, Kenneth & Keyser, Stephen Jay (eds.), *The View from Building 20*, 111–176. Cambridge, MA: MIT Press.

Haspelmath, Martin & Sims, Andrea. 2010. *Understanding Morphology* (2nd edition). London: Hodder Education. DOI: https://doi.org/10.4324/9780203776506

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic inquiry* 20(2). 253–306.

He, Wei. 1982. Huojia fangyan yunmu de fenlei [Classification of Huojia finals]. *Fangyan* [Dialect] (1). 22–36.

Hockett, Charles. 1954. Two models of grammatical description. *WORD* 10(2–3). 210–234. DOI: https://doi.org/10.1080/00437956.1954.11659524

Horwood, Graham. 2001. Anti-faithfulness and subtractive morphology. ROA 466.

Hou, Jingyi & Wen, Duanzheng. 1993. *Shanxi Fangyan Diaocha Yanjiu Baogao* [Survey of dialects in Shanxi Province]. Taiyuan: Joint Press of Universities in Shanxi.

Hyman, Larry. 1984. On the weightlessness of syllable onsets. In *Annual Meeting of the Berkeley Linguistics Society* 10. 1–14. DOI: https://doi.org/10.3765/bls.v10i0.3312

Hyman, Larry. 1985. *A theory of phonological weight*. Dordrecht: Foris. DOI: https://doi.org/10.1515/9783110854794

Inkelas, Sharon & Zoll, Cheryl. 2005. *Reduplication: Doubling in morphology*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511627712

Itô, Junko & Mester, Armin. 1998. Markedness and word structure: OCP effects in Japanese. ROA 255.

Jackendoff, Ray & Audring, Jenny. 2018. Relational morphology in the parallel architecture. In Audring, Jenny & Masini, Francesca (eds.), *The Oxford handbook of morphological theory*, 390–408. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/OXFORDHB%2F9780199668984.013.33

Kapatsinski, Vsevolod. 2017. Learning a subtractive morphological system: Statistics and representations. In LaMendola, Maria & Scott, Jennifer (eds.), *Proceedings of the 41st annual Boston University Conference on Language Development*, 357–372. Somerville, MA: Cascadilla Press.

Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17. 351–366. DOI: https://doi.org/10.1515/tlir.2000.17.2-4.351

Kiparsky, Paul. 2003. Finnish noun inflection. In Nelson, Diane & Manninen, Satu (eds.), *Generative approaches to Finnic and Saami linguistics*, 109–161. Stanford: CSLI Publications.

Kiparsky, Paul. 2010. Reduplication in stratal OT. In Uechi, Linda & Wee, Lian Hee (eds.), *Reality exploration and discovery: Pattern interaction in language & life*, 125–142. Stanford: CSLI Publications.

Kiparsky, Paul. 2015. Stratal OT: a synopsis and FAQs. In Hsiao, Yuchau E. & Wee, Lian-Hee (eds.), *Capturing phonological shades within and across languages*, 2–44. Newcastle upon Tyne: Cambridge Scholars Publishing.

Kirchner, Robert. 1996. Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27(2). 341–350. DOI: https://doi.org/10.1002/9780470756171.ch21

Köhnlein, Björn. 2018. A morpheme-based approach to subtractive pluralization in German dialects. *Phonology* 35(4). 617–647. DOI: https://doi.org/10.1017/S0952675718000222

Kurisu, Kazutaka. 2001. *The phonology of morpheme realization*. Santa Cruz, CA: University of California dissertation. DOI: https://doi.org/doi:10.7282/T3R49PKM

Kurisu, Kazutaka. 2020. The nature of subtractive processes in morphology. In *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acrefore/9780199384655.013.596

Lee, Wai-Sum & Zee, Eric. 2017. Apical vowels. In Sybesma, Rint & Behr, Wolfgang & Gu, Yueguo & Handel, Zev & Huang, C.-T. James & Myers, James (eds.), *Encyclopedia of Chinese language and linguistics (Volume 1)*, 169–172. Leiden: Brill. DOI: https://doi.org/10.1163/2210-7363_ecll_COM_00000018

Li, Charles N. & Thompson, Sandra A. 1981. *Mandarin Chinese: A functional reference grammar*. University of California Press. DOI: https://doi.org/10.1525/9780520352858

Lieber, Rochelle. 1992. *Deconstructing morphology: Word formation in syntactic theory*. Chicago: University of Chicago Press.

Lin, Yen-Hwei. 1989. *Autosegmental treatment of segmental processes in Chinese phonology*. Austin, TX: University of Texas dissertation.

Lin, Yen-Hwei. 1993. Degenerate affixes and templatic constraints: Rime change in Chinese. *Language* 69(4). 649–682. DOI: https://doi.org/10.2307/416882

Lin, Yen-Hwei. 2001. Toward a unified account of three classes of Huojia affixed words. In Xu, Debao (ed.), *Chinese phonology in generative grammar*, 223–240. San Diego: Academic Press. DOI: https://doi.org/10.1016/B978-012767670-8/50010-7

Lin, Yen-Hwei. 2004. Chinese affixal phonology: Some analytical and theoretical issues. *Language and Linguistics* 5(4). 1019–1046.

Lin, Yen-Hwei. 2010. Unexpected morphophonological outputs. In *NACCL-22 & IACL-18: Proceedings of the 22nd North American Conference on Chinese Linguistics and the 18th International Conference on Chinese Linguistics*, 363–382.

Lin, Yen-Hwei. 2015. OT factorial typology and the typology of Mandarin mid vowel assimilation. *Language and Linguistics* 16(6). 769–789. DOI: https://doi.org/10.1177/1606822X15602600

Lin, Yen-Hwei. 2022. The morphophonology of Chinese affixation. In Huang, Chu-Ren & Lin, Yen-Hwei Lin & Chen, I-Hsuan & Hsu, Yu-Yin (eds.), *The Cambridge handbook of Chinese linguistics*, 223-244. Cambridge University Press. DOI: https://doi.org/10.1017/9781108329019.014

Łubowicz, Anna. 2002. Derived environment effects in Optimality Theory. *Lingua* 112. 243–280. DOI: https://doi.org/10.1016/S0024-3841(01)00043-2

Lyu, Zhenjia. 1991. *Yuncheng fangyanzhi* [*Yuncheng Dialect*]. Taiyuan: Joint Press of Universities in Shanxi.

Manova, Stela. 2011. *Understanding morphological rules: With special emphasis on conversion and subtraction in Bulgarian, Russian and Serbo-Croatian*. Dordrecht: Springer Netherlands. DOI: https://doi.org/10.1007/978-90-481-9547-3

Manova, Stela. 2016. Subtractive morphology. In *Oxford bibliographies*. Oxford University Press. DOI: https://doi.org/10.1093/obo/9780199772810-0116

Manova, Stela. 2020. Subtraction in morphology. In *Oxford research encyclopedia of linguistics*. Oxford University Press. DOI: https://doi.org/10.1093/acrefore/9780199384655.013.572

Marantz, Alec. 1982. Re Reduplication. *Linguistic Inquiry* 13(3). 435–482.

McCarthy, John & Kimper, Wendell & Mullin, Kevin. 2012. Reduplication in Harmonic Serialism. *Morphology* 22(2). 173–232. DOI: https://doi.org/10.1007/s11525-012-9203-3

McCarthy, John & Prince, Alan. 1986/1996. *Prosodic Morphology*. New Brunswick: Rutgers University Center for Cognitive Science.

McCarthy, John & Prince, Alan. 1994. The emergence of the unmarked: Optimality in Prosodic Morphology. Amherst, MA: GLSA. DOI: https://doi.org/10.7282/t3z03663

McCarthy, John & Prince, Alan. 1995. Faithfulness and reduplicative identity. In Beckman, Jill & Dickey, Laura Walsh & Urbanczyk, Suzanne (eds.), *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, 249–384. Amherst, MA: GLSA.

Morén, Bruce. 2001. *Distinctiveness, coercion and sonority: a unified theory of weight*. New York: Routledge. DOI: https://doi.org/10.4324/9781315023779

Orgun, Cemil Orhan. 1996. Sign-Based Morphology and Phonology with special attention to Optimality Theory. Berkeley, CA: University of California dissertation.

Packard, Jerome. 2015. Morphology. In Wang, William S.-Y. & Sun, Chaofen (eds.), *The Oxford handbook of Chinese linguistics*. Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199856336.013.0045

Paschen, Ludger. 2018. *The interaction of reduplication and segmental mutation: A phonological account.* University of Leipzig dissertation.

Prince, Alan & Smolensky, Paul. 1993/2004. *Optimality Theory: Constraint interaction in Generative Grammar*. Malden, MA: Blackwell. DOI: https://doi.org/10.1002/9780470759400

Saba Kirchner, Jesse. 2010. *Minimal Reduplication.* Santa Cruz, CA: University of California dissertation.

Saba Kirchner, Jesse. 2013. Minimal Reduplication and reduplicative exponence. *Morphology* 23(2). 227–243. DOI: https://doi.org/10.1007/s11525-013-9225-5

Satoh, Akira & Feng, Liangzhen. 1989. Sansei kakuken Hōgen no bunhaku idoku [Literal and colloquial readings in Huozhou dialect of Shanxi Province]. *Chuugoku Gogaku* 236. 24–34.

Selkirk, Elisabeth. 1982. The syntax of words. Cambridge, MA: MIT Press.

Selkirk, Elisabeth. 1995. The prosodic structure of function words. In Beckman, Jill N. & Dickey, Laura Walsh & Urbanczyk, Suzanne (eds.), *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, 439–470. Amherst, MA: GLSA.

Shen, Li & Feng, Liangzhen & Tsumura, Hiroomi. 2010. Yong GIS fenxi Shanxi Huozhou fangyan yuanyin hexielyu de shuaibian xianxiang [The decline of vowel harmony in Huozhou dialect, Shanxi Province: the perspective of GIS]. *Zhongguo Yuyanxuebao* [Journal of Chinese Linguistics] 14. 137–157.

Shen, Ming. 2003. Shanxi fangyan de xiaocheng [Diminutives in Shanxi dialects]. *Fangyan* [Dialect] 4. 335–351.

Shen, Ming. 2008. Shanxi Sheng de hanyu fangyan [Dialects in Shanxi Province]. *Fangyan* [Dialect] 4. 350–360.

Siddiqi, Daniel. 2018. Distributed Morphology. In Audring, Jenny & Masini, Francesca (eds.), *The Oxford handbook of morphological theory*, 143–165. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb%2F9780199668984.013.15

Smolensky, Paul. 1993. Harmony, markedness, and phonological activity. Paper presented at *the First Rutgers Optimality Workshop (ROW 1)*. Rutgers University, New Brunswick, New Jersey.

Smolensky, Paul. 2006. Optimality in phonology II: Harmonic completeness, local constraint conjunction, and feature domain markedness. In Smolensky, Paul & Legendre, Géraldine (eds.), *The harmonic mind: From neural computation to Optimality-Theoretic grammar,* Vol II, 27–160. Cambridge, MA: MIT Press.

Stonham, John. 1994. *Combinatorial Morphology*. Amsterdam, Philadelphia: John Benjamins. DOI: https://doi.org/10.1075/cilt.120

Tian, Juan. 2009. *Shanxi Huozhou fangyan yuyin yanjiu* [Phonology of Shanxi Huozhou Dialect]. Xi'an: Shaanxi Normal University MA thesis.

Tian, Xicheng. 1986. Shanxi Heshun fangyan de zibian yunmu [Zi-changed finals in the Heshun dialect of Shanxi Province]. *Zhongguo Yuwen*, 371–373.

Tian, Xicheng. 1992. Huozhou fangyan de xiaocheng bianyun [Diminutive rime change of Huozhou Dialect]. *Journal of Shanxi University* 1. 1–8.

Topintzi, Nina. 2008. Weight polarity in Ancient Greek and other languages. In *The Proceedings of the 8th International Conference on Greek Linguistics*, 503–517.

Trommer, Jochen. 2011. *Phonological aspects of Western Nilotic mutation morphology*. University of Leipzig Habilitationsschrift.

Trommer, Jochen & Zimmermann, Eva. 2014. Generalised mora affixation and quantity-manipulating morphology. *Phonology* 31(3). 463–510. DOI: https://doi.org/10.1017/S0952675714000220

van de Weijer, Jeroen & Zhang, Jisheng. 2008. An X-bar approach to the syllable structure of Mandarin. *Lingua* 118(9). 1416–1428. DOI: https://doi.org/10.1016/j.lingua.2007.09.006

Walker, Rachel. 2001. Positional markedness in vowel harmony. In Féry, Caroline & Dubach, Antony Green & van de Vijver, Ruben (eds.), *Proceedings of HILP* 5. 212–232.

Walker, Rachel. 2005. Weak triggers in vowel harmony. *Natural Language & Linguistic Theory* 23(4). 917–989. DOI: https://doi.org/10.1007/s11049-004-4562-z

Walker, Rachel. 2011. *Vowel patterns in language*. New York: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511973710

Wang, Hongjun. 2008. *Hanyu feixianxing yinxixue* [*Nonlinear Phonology of Chinese (2nd ed.)*]. Peking University Press.

Wang, Linhui. 1993. Linyi fangyan de *zi*-wei yu *zi*-bian yunmu [The *zi* affix and *zi* rime change in Linyi dialect]. *Journal of Shanxi Normal University* 20(1). 96–99.

Wolf, Matthew. 2007. For an autosegmental theory of mutation. *ROA* 754. DOI: https://doi.org/10.7282/T3CZ3B3S

Wu, Fei & Kenstowicz, Michael. 2015. Duration reflexes of syllable structure in Mandarin. *Lingua* 164. 87–99. DOI: https://doi.org/10.1016/j.lingua.2015.06.010

Zhang, Jie. 2000. Non-contrastive features and categorical patterning in Chinese diminutive suffixation: MAX[F] or IDENT[F]? *Phonology* 17(03). 427–478. DOI: https://doi.org/10.1017/S0952675701003979

Zhang, Jisheng. 2023a. Chinese syllable structure. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. DOI: https://doi.org/10.1093/acrefore/9780199384655.013.898

Zhang, Jisheng. 2023b. Lun hanyu yuanyin yinwei [On the vowel phonemes of Chinese]. *Dangdai Yuyanxue* [Contemporary Linguistics] 1. 144–158.

Zimmermann, Eva. 2013. Non-concatenative allomorphy is generalized prosodic affixation: The case of Upriver Halkomelem. *Lingua* 134. 1–26. DOI: https://doi.org/10.1016/j.lingua.2013.06.001

Zimmermann, Eva. 2017. *Morphological length and prosodically defective morphemes*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198747321.001.0001