Reduplication and the structure of nouns in Xining Chinese

Qi Wang¹ · Anders Holmberg²

Received: 24 April 2019 / Accepted: 25 August 2020 / Published online: 10 September 2020
© The Author(s) 2020

Abstract In Xining Chinese, especially as used by older people, free nouns are always reduplicated, as a purely formal condition without any semantic effects. We argue that the reduplication takes place when an acategorial root is merged with a null nominal categorizer which copies the phonological matrix of the root, as an effect of a condition ruling out free monosyllabic nouns. When the condition is not independently satisfied, as in a compound or derived noun, reduplication is how the condition is met. Reduplication also occurs optionally in compounds or derived nouns. In conjunction with a minimalist theory of word formation, this will be shown to predict the distribution of reduplication in various contexts. For instance, the head of a compound can be reduplicated, but not the modifier, some affixes but not others permit reduplication of the base, non-compositional compounds do not allow reduplication, and so-called ‘bound roots’ (really, bound words) are not reduplicated. The phenomenon provides very strong evidence that simple content words are made up of an acategorial root and a categorizer which is often null, but can be overt in some languages, including Xining Chinese, where it is overt in nouns by virtue of reduplication.

Keywords Root · Merge · Categorizer · Compound · Chinese · Reduplication

✉ A. Holmberg
anders.holmberg@newcastle.ac.uk

Q. Wang
wangdpnc@126.com

¹ School of Foreign Languages and Literature, Wuhan University, Luojiashan, Wuchang Qu, Wuhan Shi, Hubei, 430072, China

² School of English Literature, Language, and Linguistics, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK

1 Introduction

A morphological peculiarity of the variety of Chinese traditionally spoken in and around Xining in the northwest of China is that common nouns are always reduplicated, as exemplified in (1).

- (1) a. Jiá sa **fēi fēi** hū liāo.
 she PRT water water drink PRF
 ‘She has drunk water.’
 b. Jiá sa **mǒ mó** ha mēi hà zhei.
 she PRT steamed.bun steamed.bun OBL buy PRF PRT
 ‘She has bought steamed buns.’
 c. Zhǐ gó **dō dō** sa hǔdú guǐ na.
 this CLF knife knife OBL very expensive PRT
 ‘This knife is very expensive.’

The reduplication has no semantic effect whatsoever, but is a purely formal requirement. In particular in the variety of Xining Chinese spoken by the older generation, which we will refer to as Traditional Xining Chinese, the reduplication is compulsory. Similar reduplication is common also in other dialects spoken in northwest China. Our data are exclusively from Traditional Xining Chinese, though, abbreviated TXC (called Old Xining Chinese in Wang 2018).¹

The reduplication, we claim, is required to satisfy a language-particular condition that a free noun must consist of at least two syllables. There is a well-known tendency in Chinese to favour disyllabic words over monosyllabic ones (Duanmu 1999). An effect of this is a strong preference for compounds of various kinds. Another effect, characteristic of TXC, is reduplication of nouns, compulsory with monosyllabic nouns, optional in certain disyllabic nouns in ways to be discussed in the paper.

The analysis we propose here is that the reduplication of nouns in TXC is the result of the copying of the phonological features of the root by a nominal categorizer. It is based on the premise that lexical categories are made up of a root devoid of a syntactic category feature, merged with a categorizer, that is a functional head encoding syntactic category (Marantz 1997; Josefsson 1997, 1998; Harley and Noyer 1999; Lieber 2006; Embick and Noyer 2007, 2008; Harley 2011, 2014; Embick 2015; Hu and Perry 2017). The categorizer is often a null morpheme. That is the case with lexical categories generally, for example in Standard Mandarin (henceforth called Mandarin), except in some cases where the category is provided by an overt affix. It is also the case in TXC with categories other than the noun. But for nouns in TXC, the nominal categorizer is mostly overt. If it is not realized as an affix, it is realized by copying the phonological features of its sister root. We will refer to this as noun reduplication, as a convenient descriptive term, but with the understanding that this is not reduplication of a noun, but reduplication of a root, by copying its phonological features onto a null nominalizer.

This hypothesis makes a number of predictions about contexts where reduplication will be found, predictions that are all met. This means that we can always tell a root

¹ One of the authors is a native speaker of TXC. The data are checked with other speakers of TXC, including speakers that are older than 70.

from a noun in Xining Chinese: a root not accompanied by a nominal categorizer will not be reduplicated, while a root merged with a nominal categorizer will be. In this way the reduplication serves as a probe into the structure of words,² particularly nouns, in TXC. This will be shown to shed new light on controversial categories in Chinese morphology, including various kinds of compounds and the category called bound roots in the literature (Packard 2000). On a more general level, reduplication in TXC provides strong evidence for the hypothesis that content words are made up of an acategorial root merged with a designated categorizer. This hypothesis is widely but not universally assumed within generative morphosyntax (see Borer 2014a for a partial rebuttal), and is even more controversial in more traditional morphological theory.

The rule which derives the reduplicated form is phonological in that it operates on a phonological unit, the phonological matrix of a root, with effects only at PF (the Articulatory-Perceptual interface), none at LF (the Conceptual-Intentional interface). It is a morpho-syntactic rule in the sense that the conditions on the reduplication are syntactically defined. First, it concerns specifically nouns. It is thus not, for example, motivated by a general condition on the size of minimal words in TXC (McCarthy and Prince 1990; Hall 1999). Verbs and adjectives are not usually reduplicated in TXC, and if they are, it has a semantic effect, denoting repetition or intensification, among other things. Second, the reduplication is a matter of copying of features by a syntactically derived head, the categorizer *n*, which furthermore can take place if and only if the root and *n* are sisters; adjacency is not sufficient. The upshot is that the reduplication of nouns in TXC is a postsyntactic rule but dependent on a specific syntactic input, categorial and structural.

The paper has the following structure: Sect. 2 provides some background information on the language. Sect. 3 outlines the basic theoretical assumptions. This will include a review of a certain kind of nominal compound in Swedish, because it provides a particularly clear case of overt root-noun distinction, to be compared with the overt root-noun distinction in TXC. In Sect. 4 we go through a set of correct predictions made by our theory concerning the distribution of noun reduplication in compounds and affixed words in TXC. In Sect. 5 we discuss a class of items characteristic of Chinese, called bound roots or stems in the literature, but that we will refer to as bound words. These cannot be reduplicated in TXC, a finding with interesting theoretical consequences discussed in this section. Sect. 6 will be about non-compositional compounds. Our theory, conforming with Zhang's (2007) and Hu and Perry's (2017) theory in relevant respects, correctly predicts the absence of reduplication in such compounds. Sect. 7 discusses recursive compounding, which presents a challenge to the theory assumed. Sect. 8 is the conclusions.

2 Background: Morphology of traditional Xining Chinese

The city of Xining is situated in the northwestern part of China and is the capital city of Qinghai province. What we call Traditional Xining Chinese (TXC) is what is

²The point that reduplication can serve as a probe into the structure of words and phrases is also made by Travis (2001). The reduplication that we describe does not fall into any of the classes of reduplication that Travis identifies, but we will return to Travis (2001) in the conclusions.

called ‘old Xining dialect’ in Dede (2006). It is a language which is spoken by the Han and the Hui people, especially the elderly, who live in the city of Xining and nearby villages, that is, Huangyuan, Huangzhong, Pingan, Menyuan, Huzhu, Guide and Hualong (Zhang and Zhu 1987:1–4; Zhang 2001:1–3; Dede 2003:334–335). Some older speakers of other ethnic groups in the area also use TXC as their second language. This paper will concern TXC as spoken by the Han.

TXC belongs to the group of Central Plains Mandarin (Dede 2003:330–331; Wang and Dede 2016:407) and is incomprehensible to (Standard) Mandarin, which is a variety of Northern Mandarin (Huang and Shi 2016:2; Po-Ching and Rimmington 2016:3). TXC has its own distinctive character, which is due to the prolonged contact with non-Sinitic languages used by other ethnic groups in the region (Dede 2006:320–321, 327). In terms of syntax, SOV is the dominant word order. There is a wider range of functional categories than in Mandarin: a case marker, for instance *-ha*, is often needed for the preverbal object (Dede 1993:72; Dede 2006:327; Bell 2017; see examples under (1)), and there is a clause-final complementizer and a VP-final tense-marker (Dede 1999:51; Bell 2017, 2019). In terms of morphology, TXC is not so different from Mandarin, though, except for the obligatory reduplication of free nouns which is characteristic of TXC (Wang 2018). Affixation and compounding are used as ways of forming nouns in both Mandarin and TXC, but for TXC, there is another way of forming nouns in a regular fashion, which is reduplication.

3 Roots and categorizers: The structure of compounds

3.1 Merge and the structure of Swedish nouns

In this section we will first outline our basic theoretical assumptions. As mentioned, a crucial assumption is that content words are made up of an acategorical (categoryless) root merged with a designated categorizer, which may be null, and often is, which is one reason why the assumption is controversial (Lieber 2006; Lehmann 2010). We will argue that the categorizer in the case of nouns is usually overt in TXC, being realized as reduplication. As a preliminary we will discuss another, already investigated case of overt nominal categorizer, in a class of Swedish nouns. The point of this comparison is to establish what we expect to see if the reduplication in TXC is another case of overt nominal categorizer.

We assume that words are composed by the same rule as phrases, that is Merge in the sense of Chomsky (1995:243) and subsequent work within the Minimalist program, following much work within Distributive Morphology: Harley and Noyer 1999; Borer 2005, 2013, 2014a; di Sciullo 2005; Embick and Noyer 2007, 2008; Harley 2011, 2014; Embick 2015.

- (2) Merge α and β to form a set $\{\alpha, \beta\}$ with a label γ , where γ is = either α or β , depending on which one is the head.

Following standard practice we represent the set as a tree. The two trees formed by α and β are (3a,b)

That α and β make up a set, rather than a pair, means that they are not linearly ordered by Merge. That is to say, order is not a matter of phrase-syntax or word-syntax, but is a morpho-phonological matter, part of the externalization of the structure built by Merge (in Chomsky's 2013 sense). We also assume that labelling of the set formed by Merge is crucial for the set to function as a morphological or syntactic unit (Hornstein 2009; Chomsky 2013). Correspondingly, the definition of 'head' we assume is: The head of a set $\{\alpha, \beta\}$ is the member that determines the label, hence the category, of the set (in Sect. 6 we will discuss sets formed by Merge which are unlabelled, showing that they have special properties). Essentially following Chomsky (2013) we take all this to mean that the grammar must have the means to determine which constituent of a set $\{\alpha, \beta\}$ is the head.

In this light, consider the structure of compounds. Compound nouns (*wallpaper*, *love affair*, *wallpaper design*) are traditionally taken to be made up of two or more nouns, forming a noun. For example *wallpaper design* would have the structure (4).

Traditionally, following early work in generative morphology (Lieber 1980; Williams 1981), the head in English words is determined by the Righthand-Head Rule (Williams 1981): the rightmost constituent of a word is the head. In the case of (4) this entails that *paper* is the head of *wallpaper* and *design* is the head of *wallpaper design*. However, if Merge builds sets without linear order, then the linear order cannot determine whether α or β in $\{\alpha, \beta\}$ is the head. Instead, something else determines whether α or β in $\{\alpha, \beta\}$ is the head, and this, in turn, determines the linear order of α and β . The 'something else' is the syntactic features of α and β and the relation between them, which must be asymmetric in such a way that one of them is the head, i.e. one of them provides the label for the set $\{\alpha, \beta\}$ constructed by Merge. We are familiar with asymmetry in syntax: phrases are made up of a word-level category merged with a maximal phrase-level category, and the word-level category is the head. Merging two phrases presents a problem for labelling, though, which has to be resolved; this is the basis for Chomsky's (2013) 'labelling algorithm,' building on Moro (2000). If words are constructed by the same operation Merge as phrases, we expect to see analogous asymmetry in words, including compounds (cf. di Sciullo 2005). While English nominal compounds do not show any obvious visible signs of asymmetry, there are languages where they do. One such language is Swedish.

Swedish has a large class of nouns (the so-called first and second declensions) which are made up of a root plus a vowel *-a* or *-e* (Kiefer 1970; Holmberg 1992; Josefsson 1997, 1998):

- (5) skol-a, flick-a, poj-k-e (Swedish)
 school-A, girl-A, boy-E

The vowel can be analysed as encoding singular number, as it alternates with plural suffixes: *skol-or* ‘schools’, *flick-or* ‘girls’, *poj-k-ar* ‘boys’. For nouns denoting people and gendered animals it also encodes semantic gender: *-a* for females, *-e* for males. The analysis of the vowel is controversial. According to Kiefer (1970) it has no features except a declension class feature, while according to Ejerhed and Bromley (1986) and Holmberg (1992) it has the number feature [-PL]. Inspired by Harris’s (1991) theory of Spanish word formation, where Harris argues that the Spanish nominal vowels *-a* and *-o* (as in *muchacha*, *muchacho*) are not gender markers but ‘word markers’ that serve to make a word out of a root, Holmberg (1992) assumes that the Swedish vowels *-a* and *-e* in (5) are word markers, whose “essential grammatical function is to combine with a Root, and thereby permit the Root to take part in phrase- and word-syntactic processes.” According to Holmberg (1992) the root is a ‘maximal category’ not projecting any features onto a dominating node, which allows the vowel to project its features. The question is taken up again in Josefsson (1997, 1998), who argues that the reason why the root does not project any categorial feature is that it does not have a categorial feature, and therefore always has to combine with an item that has one. The noun class exemplified in (5) is a special case, where a nominal feature is realized as a vowel.

The idea that roots are acategorial has since become standard within the Distributed Morphology model of morphology (Halle and Marantz 1993; Marantz 1997, 2007; Harley and Noyer 1999; Borer 2005, 2013, 2014a; di Sciullo 2005; Embick and Noyer 2007, 2008; Harley 2011, 2014; Belder 2011; Embick 2015; Hu and Perry 2017). In this model there is no single list of words as in the traditional Lexicon (Chomsky 1965). Instead, words are constructed as part of the syntactic derivation by combining items from different lists: One is the list of formatives, including the roots and the functional morphemes. Another is the list of exponents, that is vocabulary items, including those of functional categories (even particularly listing functional vocabulary, depending on the theory; see Embick 2015). Yet another is the Encyclopedia, containing the non-compositional meanings of words and idiomatic phrases. The derivation of words and phrases proceeds by merge of roots and functional categories, as described above, until the point where the derivations splits; the derived syntactic structure is ‘transferred to PF and LF’ separately; call it the point of Transfer (Chomsky 2001, 2008). On the PF-branch, the structure undergoes Vocabulary Insertion, Linearization, and other rules affecting the phonologically realized form but not the meaning. A controversial issue, within this theory, is the nature of roots. One question is whether roots are identified in the syntax by their phonological features, semantic features, or neither. Harley (2014) articulates a radical version of the model, where the roots in the list of formatives are purely abstract objects, identified only by an index which ensures that the root receives a particular phonological form at Vocabulary Insertion in the derivation of PF, and a particular semantic interpretation in the derivation of LF, after Transfer. According to a more conservative version of the theory, following Borer (2014a, 2014b) and Embick (2015), roots are identified phonologically in the syntax already. The facts we discuss and the generalizations we formulate can be modelled in either version of the theory. We will, however, represent the roots and functional categories by their phonological, written form, for ease

Another controversial issue concerning roots is when they merge with their categorizer. According to Marantz (1997) roots are inert until they are merged with a categorizer or other item with syntactic category, thus forming part of a syntactic object with a categorial label. Harley (2014), on the other hand, articulates a version of the theory where roots behave like other syntactic heads, in that they can take a complement, project a phrase (a Root Phrase), and undergo head movement to a categorizer head. We assume a ‘conservative’ theory of word formation, in line with Marantz (1997), Borer (2014a, 2014b), Embick (2015), where roots do not directly take part in syntactic operations, but do so only by way of merging with a categorizer or other category-bearing item. That is to say, we do not adopt the idea that roots take complements, project phrases, and undergo head movement.³

(6)

N

R

skol

[1]

n

-PL

~~u1~~

The root form shows up in derived words:

- The structure of *flickaktig* ‘girlish’, for example, is (8), the acategorial root *flick* merging with the adjectival suffix *-aktig*, projecting an adjective.

⁵ A slightly more complex analysis may be considered, where the root first merges with [n, u1], deriving a noun, which subsequently merges with the number feature. The so-derived structure [_N [_N R [n u1]] –PL] would be spelled out as *skola* (in the case at hand).

The form *flicka-aktig* in (7), merging a noun *flicka* with an adjectival suffix, is ill-formed on account of merging two items each projecting a categorial feature, meaning, we assume, that neither projects, so that the resulting construct does not have a category (Josefsson 1997, 1998).

Another place where the root form shows up is as modifier of attributive compounds:⁶

- (9) skol-flicka, flick-skola, *skola-flicka, *flicka-skola, *skol-flick, *flick-skol
 school girl, girl school (Swedish)

The structure of for example *skolflicka* ‘schoolgirl’ would be (10) (representing the nominalizer feature complex as ‘n’)

Here, the root *flick* merges with the nominalizer spelled out as *-a*, projecting a noun.⁷ This noun merges with the root *skol*. Since the root has no categorial feature, the noun labels the dominating node, deriving a compound noun denoting a kind of girl. The ungrammatical form **skola-flicka* is ruled out as being ‘too symmetrical’: It would be composed of two merged nouns, with the grammar unable to determine the head. The form **skolflick* is ruled out as well, being composed of two merged roots, which yields a construct without categorial or other syntactic features.

By hypothesis, the derived noun has no linear order until it is sent to PF. In this process the linearization rule corresponding to Williams’s (1981) Righthand Head Rule applies to (6), (8), and (10), deriving the strings *skola*, *flickaktig*, and *skolflicka*. The rule for Swedish can be stated as:

- (11) A set $\{\alpha, \beta\}$ where α is the head projecting a word is linearized as $\beta > \alpha$.

⁶Here and throughout the article we use the term *attributive compound* for endocentric compounds consisting of a head and a modifying/restricting non-head, including compounds where there is, arguably, a thematic relation between the head and the non-head, as in *toothbrush*, *snow shovel*, etc.

⁷The fact that merge of a root with a nominalizer is sensitive to declension class (see (6)) prompts a reviewer to ask whether merge of a root with a word, as in (10), is sensitive to declension class. The answer is no. This is ensured by the feature analysis proposed: The root has no unvalued features, thus places no restrictions on merge.

As discussed, following the trend in Distributed Morphology, we assume that common nouns generally have the structure $[_N R, n]$, not just in Swedish but universally (although we shall also claim that exceptions exist, which have particular, identifiable properties). In many cases, perhaps most cases across languages, the categorizer is realized as null.

The condition that the non-head of a compound must be a root cannot hold universally, in view of left-recursive compounding, as in *football pitch*, *high-street fashion*, etc. We will return to this issue in Sect. 7. The issue is avoided in TXC, though, as this variety of Chinese does not allow recursive compounding.

3.2 The structure of common nouns in Traditional Xining Chinese

Free common nouns in TXC are always reduplicated; see examples in (1). As in other languages, a common noun in TXC is minimally made up of a root and a nominalizer *n*, merged in the syntax. The nominalizer *n*, like its verbal and adjectival counterpart, can only merge with an acategorial unit, that is typically with a root, as dictated by the morphological implementation of Chomsky's (2013) labelling algorithm: a set consisting of a categorized item, say a verb, and a categorizer *n* could not be labelled, being 'too symmetrical'.⁸ As other functional categories, and following the Distributed Morphology model, the nominalizer is provided with phonological form at Vocabulary Insertion, in the derivation of PF after the Transfer point. In many languages, including Mandarin, the nominalizer *n* is mostly realized as phonologically null. However, as discussed above, with many Swedish nouns, *n* is realized as *-a* or *-e*. And in the case of TXC, *n* is mostly realized as a phonological copy of its sister root. At Transfer, the structure of, for example, the noun *sū sǔ* 'lock' is (12) (assuming that roots have their phonological matrix in the syntax already). At Vocabulary Insertion the phonological matrix of the root, $[sū]$, is copied by *n*, and the word is spelled out *sū sǔ*.⁹

$$(12) \quad \begin{array}{c} N \\ \swarrow \quad \searrow \\ R \quad n \\ | \quad \quad \rightarrow \quad sū \quad sǔ \\ sū \end{array} \quad (TXC)$$

This is a phonological feature copying operation, but operating under a syntactically defined condition: The root and *n* have to be sisters.¹⁰

⁸The reason for the formulation "typically a root" will become evident in Sect. 6, where we discuss cases of derived complex roots, which are acategorial units mergeable with *n*.

⁹The root of the noun *sū sǔ* 'lock' also occurs as a verb meaning 'lock,' predictably without reduplication.

(i) Nī bǎ mén sū ha liao.
you BA door lock PRT PRF
'You have locked the door.'

¹⁰The effect of noun reduplication on tone is an interesting issue which we leave for future research. The tone is often reduplicated (see examples (1c), (27b), (29b)), often not (see (1b), (12), (15)). There are no obvious generalizations that stand out when more data are taken into account. There are also tone sandhi effects which we will not attempt to characterize, but leave for future research. See Zhang and Zhu (1987) for a brief discussion of tone sandhi in Xining Chinese (not mentioning reduplication).

As for the formal trigger of the reduplication, it is not the case that *n* triggers the operation, say, by virtue of an unvalued feature (Chomsky 2001 *et seq.*) or because it is somehow ‘defective’ (Harley 2004).¹¹ The nominalizer *n* does not ‘need’ phonological features. This will be seen in Sect. 6, where we discuss cases where reduplication does not take place because the sisterhood condition is not met, but where the so derived nouns are nevertheless well-formed. We propose, instead, that the reduplication is an optional phonological feature copying operation allowed by *n*, but forced in a case like (12) by a language-particular output condition or ‘filter,’ applying after Vocabulary Insertion in TXC.

(13) **N* if *N* is a free word and has less than two syllables.

If the free noun is a compound, reduplication need not apply, as (13) is satisfied anyway. But if the structural condition is met, that is if the compound contains a root with a sister nominalizer, reduplication may take place. In Sect. 4 the optionality of noun reduplication will be exemplified by various complex nouns which independently have two syllables, and where reduplication can take place or not.

The formulation (13) also implies that nouns with a polysyllabic root need not undergo reduplication. Now, Chinese is notorious for its close correspondence between syllable and morpheme, and, in addition, grapheme; Norman (1988), Basciano and Ceccagno (2009), Chung et al. (2014). In the words of Norman (1988:154): “[...] the great majority of morphemes coincide phonologically with a monosyllable; stated conversely, almost every syllable can be analysed as an independent morpheme.” That is to say, roots are generally monosyllabic. There are loanwords, though, which are polysyllabic. An example in TXC is the word *hǎdá*, referring to a piece of silk used as a greeting gift among Tibetan and Mongolian people. (13) predicts that this noun does not need to undergo reduplication, a correct prediction. We illustrate this by comparing *hǎdá* with the monosyllabic noun *mó* ‘steamed bun’.

(14) *bǐ* *zháozí* *hǎdá* (**hǎdá*) *yōu* *mò*?
 white PRT HADA have Q
 ‘Do you have a white hada?’

(15) *mǒ* *mó* *ha* *chī* *bāo*?
 steamed.bun PRT eat Q
 ‘Do you want steamed buns?’

The fact that reduplication in (14) is not just unnecessary but ill-formed is an indication of another condition on the reduplication: Only a single syllable can be redupli-

¹¹ A reviewer for *NLLT* suggests that the copying can be regarded as a variety of Conflation, a mechanism suggested but ultimately rejected by Hale and Keyser (2002) but revived by Harley (2004). Simplifying somewhat, Conflation is movement of the phonological features of a complement to a head, triggered when the head is somehow phonologically defective. Phonological feature copying, the reviewer suggests, would be a variety of Conflation where the original copy is not deleted. However, as will be discussed in Sect. 6, *n* in TXC can be spelled out as null when copying is ruled out. That is to say, it is not defective.

cated.¹² We will see other effects of this condition in Sect. 6.¹³

Our formulation of noun reduplication in TXC entails that it applies to ‘lexical nouns’ only, that is nouns made up of a root and a nominalizer. Pronouns do not undergo reduplication, nor do nominal quantifiers.¹⁴

(16) nǎo (*nǎo) ‘I’, nī (*nī) ‘you’, jiá (*jiá) ‘he/she’, fěi (*fěi) ‘who’ (TXC)

Proper names are an interesting case. They can be reduplicated, and often are, as pet names. This is true not just in TXC but in Mandarin as well (and other varieties of Chinese). This seems to be a different kind of reduplication, though, than in (12), indicated by the fact that it can have a semantic effect, namely endearment. This is true of TXC as well as Mandarin. We will, however, leave proper names aside in the present paper.

Like other varieties of Chinese, TXC also has reduplication of verbs and adjectives (cf. Ren 2006; Wang 2009). This is clearly a different phenomenon, though, from the nominal reduplication we are investigating here. It is always associated with a semantic effect: with verbs repetition or continuation, and with adjectives intensification. It is thus not obligatory. We take this to mean that the reduplication does not apply to roots, but to words; that is, roots with either a verbal or an adjectival categorizer, affecting their interpretation as verbs and adjectives. We are sympathetic to the idea, articulated by Travis (2001), that the reduplication is a matter of phonological feature copying in this case, too, but by a head encoding, for instance, intensity, merged with an adjective, or perhaps an AP. We will not deal with this kind of reduplication in this paper, though.

¹²While this appears to be common as a feature of reduplication, there are many cases of disyllabic reduplication reported in the literature; see e.g. Frampton (2009:117–128).

¹³Reduplication also cannot apply to a part of the polysyllabic root *hǎdá*: **hǎhǎdá*, **hǎdǎdá*. This follows from the condition that noun reduplication applies only to a root merged with a null categorizer.

¹⁴A widely assumed analysis of pronouns is the one pioneered by Postal (1969) according to which the spelled-out pronoun is the head of a DP with a complement NP which is mostly null. As a functional head, the pronominal D is predicted not to reduplicate in TXC. Arguably the most convincing argument in favour of Postal’s analysis is the fact that pronouns, in particular plural pronouns, can occur with an overt NP complement, as in *we teachers*, *you children* (see Höhn 2017). This construction is found in TXC as well:

(i)	nī-men xiù sǎng,	nǎo-men jiǒu yuān
	2 -PL school student	1 -PL education member
	‘you students’	‘we teachers’

As predicted, if the noun is monosyllabic (is not a compound), it will be reduplicated:

(ii)	nǎo-men mǎo mó
	1 -PL cat cat
	‘we cats’

crucial for the reduplication to take place in TXC nouns. Based on this and the classification of the object-denoting item *xiǒng* ‘countryside’ and *róu* ‘meat’, it is predicted that items such as *xiǒng* ‘countryside’ and *róu* ‘meat’ which are merged with the suffixes *-bóng* and *-dán*, as in (17), cannot be reduplicated, in the absence of a sister null nominalizer which would copy their phonological features. This prediction is true, as shown by the following examples:¹⁵

(19) a. **xiǒng xiǒng-bóng* (TXC)
countryside countryside-person

b. **rǒu róu-dán*
meat meat-person

As independent nouns, *xiǒng* ‘countryside’ and *róu* ‘meat’ can and must be reduplicated. In this case they are merged with the semantically and phonologically null nominalizer, which will copy their phonological features in order to comply with the two-syllable condition (13). We will from now on gloss the reduplicant as ‘n,’ reflecting the analysis that it is the categorizer spelled out as a copy of the sister root.

(20) a. *xiǒng xiǒng* (TXC)
countryside n
‘countryside’

b. *rǒu rǒu*
meat n
‘meat’

(19a,b) are not ruled out by a condition which excludes reduplication in derived words, or excludes reduplication in words which already have two syllables. As we shall see directly, reduplication does occur in derived words, resulting in words with at least three syllables.

The suffixes *-bóng* and *-dán* themselves cannot be reduplicated, either, neither on their own, nor when merged with a root.

(21) **bǒng-bǒng*, **dǎn-dǎn*, **xiǒng-bǒng-bǒng*, **róu-dǎn-dǎn* (TXC)

This follows if they are pure functional heads, not made up of a root and a categorizer. They are the spell-out of a small bundle of syntactico-semantic features and a nominal feature. They are members of a list of nominalizers which also includes the semantically vacuous nominalizer.

This analysis of *-bóng* and *-dán* may be called into question. Josefsson (1997, 1998) argued that certain derivational affixes in Swedish are, in fact, roots. More recently the analysis of derivational affixes has been debated again; see de Belder (2011), Lowenstamm (2015), Creemers et al. (2018). Lowenstamm (2015), for example, argues that the derived noun *librarian* has basically the following structure, using our notation.

¹⁵As will be demonstrated in Sects. 4.1.2 and 4.2 there is no general ban against reduplication leading to trisyllabic words.

The suffix appears to have no effect on either the semantics or the category of the word: *mǒ-é* is a noun which means ‘cat’ and *zā-zī* a noun which means ‘powder.’ Alternative forms are the reduplicated forms *mǒ mó* ‘cat’ and *zā zā* ‘powder’ (reduplication needed to satisfy the filter (13)). A possible analysis is that *-é* and *-zī* are exponents of *n*, alternatives to reduplication. However, the suffix *-é* can occur with adjectives as well, as an optional addition, so it is not an exponent of *n*:

- (25) a. *jiēng(-e)* (TXC)
clever-E
‘clever’
b. *zhān(-e)*
flat-E
‘flat’

As in (24), the addition of *-é* has no effect on the semantics of the adjective. It is optional since adjectives can have a null categorizer. We take it to be an item devoid of any features other than an index linking it to a phonological matrix, and a selection feature selecting merge with a noun or adjective. We take this to be true of *-zī* as well, except it can only merge with a noun. If so, the other constituent in (24a,b), *mǒ* and *zā*, must be a noun, and the other constituent in (25a,b) must be an adjective, providing a head for the word. It cannot be a bare root, or the word would have no category. By hypothesis, in (24a,b) this means that it is made up of a root and a nominalizer. The structure of for example (24a), after Vocabulary Insertion, inserting the null exponent of *n*, would be (26):

A prediction can be made based on this analysis of *mǒ* ‘cat’ and *zā* ‘powder’ in (24a,b) and the procedure of reduplication in TXC nouns described in Sect. 3, which is that in the resultant affixed word, the item that the non-head suffix is merged with, can be reduplicated. This prediction is borne out, as shown in the following examples:

- (27) a. *mǒ mó-e* (TXC)
cat *n* -E
‘cat’
b. *zā zā-zī*
powder *n* -ZI
‘powder’

Comparison of (24) and (27) shows that the reduplication is optional. This, we contend, is because filter (13) is satisfied in (24) already without reduplication, by the suffix. This means, though, that the reduplication is not strictly a last resort operation. Where the conditions for the operation are met, that is where there is a root and sister null nominalizer, the reduplication may apply. If (13) is not otherwise met,

the reduplication must apply. This entails, as mentioned in Sect. 3, that the nominalizer *n* in TXC does not itself need reduplication; it is not, for example, in any way defective.^{16,17}

We also have prefixes in TXC which do not contribute the category or the semantics of the resultant affixed word:

- (28) a. ā-yī (TXC)
 A-grandfather
 ‘grandfather’
 b. gá-chēi
 GA-bike
 ‘bike’

The addition of the prefix is optional. An alternative to (28a,b) is the reduplicated forms *yǐ yǐ* and *chēi chēi* (reduplication needed to satisfy (13)). The affix makes no contribution to the interpretation of the words. We take this to mean that the category that the prefixes merge with is a noun, which is the head of the affixed word, exhaustively determining its syntactic and semantic features. If the item the prefix is merged with is a noun, it is made up of a root and a nominalizer. This predicts, given that reduplication is optional, that it can be reduplicated, a correct prediction.

- (29) a. ā-yǐ yi (TXC)
 A-grandfather n
 ‘grandfather’
 b. gá-chēi chēi
 GA-bike n
 ‘bike’

The structure of, for example (28b) would be as follows, after Vocabulary Insertion.

- (30)
-
- ```

 N
 / \
 gá N
 / \
 R n
 / \
 chēi ∅

```

<sup>16</sup>A reviewer points out, correctly, that the notion of a syntactic category, such as *-e* and *-zi*, with only phonological features seems paradoxical. However, the fact that they select certain syntactic, not phonologically defined categories implies that they are syntactic categories. Furthermore, the fact that they can be found in reduplicated words shows that they are not triggered in order to satisfy a phonological condition, in the manner of epenthetic vowels. They are items which are available, apparently, in the list of morphemes, and which can be employed to satisfy phonological conditions, if need be, but can also occur redundantly.

<sup>17</sup>Consideration of the options *mǒ mó*, *mǒ-é*, and *mǒ mó-é* indicates that the filter (13) steps in only once the whole word has gone through Vocabulary Insertion. Assuming the conventional from-root-outwards application of Vocabulary Insertion (Bobaljik, 2000), if (13) would apply at every step of Vocabulary Insertion, the word *mǒ-é* would not be derivable, as it would always be bled by *mǒ mó*. Thanks to a reviewer for *NLLT* for pointing this out.



Why does TXC have suffixes which do nothing except add a syllable to an otherwise monosyllabic word? Although it has the welcome effect of satisfying (13), this is not the (only) reason why they are employed, as (a) adjectives can take such an affix as well, and (b) Mandarin has counterparts of these affixes, even though (13) is not in force in this variety of Chinese (cf. Wang 2018). There is a well-known, strong inclination in Chinese to favour disyllabic words over monosyllabic ones. According to a comprehensive survey from 1959 (ZWGW 1959), cited in Duanmu (1999), only 27% of the 3000 most commonly used words in Standard Chinese were monosyllables. According to a more recent study of new words in Chinese (Li and Bai 1987, cited in Duanmu 1999), not a single new word is monosyllabic. Given the close correspondence in Chinese between syllable and morpheme (Norman 1988:154; Basciano and Ceccagno 2009; Chung et al. 2014) most of the disyllabic words, by far, are compounds (as will be discussed more in Sect. 6), but employing purely phonological suffixes is another way to avoid monosyllables. In this light, the reduplication in TXC is yet another, particularly radical, way to avoid monosyllables.

(31) duǎng-mēn (TXC)  
hole -PL  
'holes'

(32) duǎng duēng-men (TXC)  
hole n -PL  
'holes'

(33)

```

 N
 / \
 N -mēn
 / \
 R n
 / \
duǎng Ø

```

Springer



compounds, we predict that the head can be reduplicated, while the non-head cannot, another correct prediction. In short, the distribution of the reduplication in TXC nouns is predicted on the basis of morphosyntactic properties of the components of the nouns: If the item is lexical (root-based) and is the head of the noun, it can reduplicate, if it is not the head or is not lexical, it cannot reduplicate. Whether a word component is head or not is not phonologically marked in TXC. Thus a purely phonological approach which does not take word-syntactic structure and category into account cannot explain the distribution of reduplication in TXC nouns.<sup>19</sup>

## 5 Bound words

Chinese has a very large class of content words distinguished by the property that they have to be morphologically bound. In this section we will consider the observation that these items cannot undergo reduplication in TXC, as noted in Wang (2018), and discuss the theoretical consequences that this observation has.

(37) shows some examples of such items from Mandarin and TXC, exemplifying nouns, adjectives, and verbs.

- (37) Mandarin: *nǎo* ‘brain’, *diǎn* ‘dictionary’, *yǐng* ‘film’, *guì* ‘rule’, *shí* ‘stone’, *shé* ‘tongue’, *yǐ* ‘chair’, *lì* ‘beautiful’, *piào* ‘pretty’, *xíng* ‘walk’, *yùn* ‘carry’  
 TXC: *yī* ‘clothing’, *tā* ‘inner shirt’, *cī* ‘drawer’, *bí* ‘duvet’, *gēi* ‘armpit’, *jiéng* ‘towel’, *nóng* ‘dirty’, *jī* ‘solid’, *duèng* ‘to nap’, *jiāo* ‘disturb’

None of these items can stand alone as a free word in a phrase or a sentence (CLF = classifier; the offending word is in boldface).

- (38) a. \**yí gè shé* (Mandarin)  
           a CLF tongue  
           Intended reading: ‘a tongue’  
       b. \**zhǐ gó bí* (TXC)  
           this CLF duvet  
           Intended reading: ‘this duvet’  
       c. \**nǐ xūyào yōu xì tǒng*<sup>20</sup> (Mandarin)  
           you need excellent system interconnected system  
           Intended reading: ‘You need to make the system excellent.’

<sup>19</sup> A reviewer sketches an alternative proposal for noun reduplication, which (a) assumes that reduplication applies in order to avoid monosyllabic nouns (as in the present theory), and (b) allows the combination of affixes with reduplicated forms. In this alternative, allegedly all that needs to be said to explain the range of facts discussed in Sect. 4 is that once a noun has two syllables, it cannot be reduplicated anymore. This does indeed account for why the derived words *xiǒng-bóng* and *róu-dán* in (17) cannot reduplicate. It does not, however, account for the facts about optional reduplication discussed in Sect. 4 without various auxiliary assumptions regarding the order of application of reduplication and affixation, and some condition(s) on reduplication and/or compounding to rule out *měi měi hú* but allow *měi hǔ hú* (see (34)–(36)). In the present theory this is because the form *měi měi* can only be a noun, not a root, but a noun cannot be the non-head of compound in the languages we discuss. This and all the observations discussed in this section follow with no assumptions over and above those set out in Sect. 3.

<sup>20</sup> *xì tǒng* is a so-called synonymy compound, to be discussed below in Sect. 6.2.2.

- d. \**Jiǎ zī fǒng fǒng hǔdú kuān* na. (TXC)  
 he ZI house n very spacious PRT  
 Intended reading: ‘His house is very spacious.’

(38a,b) show that *shé* ‘tongue’ and *bí* ‘duvet’ cannot be free words in a noun phrase, while (38c,d) show that the adjectives *yōu* ‘excellent’ and *kuān* ‘spacious’ cannot stand alone as free words in a noun phrase or sentence, respectively. Instead, this type of item can only occur as part of a word, either an affixed word, as in (39a,c), or a compound, as in (39b,d). This is true of all of the items in (37).

- (39) a. *yí gè shé-tòu* (Mandarin)  
 a CLF tongue-TOU  
 ‘a tongue’  
 b. *zhǐ gó miǎn bí* (TXC)  
 this CLF cotton duvet  
 ‘this cotton duvet’  
 c. *nǐ xūyào yōu-huà* xì tǒng (Mandarin)  
 you need excellent-HUA system system  
 ‘You need to make the system excellent.’  
 d. *Jiǎ zī fǒng fǒng hǔdú kuān zhān*<sup>21</sup> na. (TXC)  
 he ZI house n very spacious flat PRT  
 ‘His house is very spacious.’

(40) shows that a bound word cannot be reduplicated in TXC. Reduplication, apparently, does not satisfy the boundness condition.

- (40) a. \**bǐ bǐ* hòu zhao hēn na. (TXC)  
 duvet n thick PRT very PRT  
 Intended: ‘The duvet is very thick.’  
 b. \**guǎi zī tā tā*  
 expensive ZI inner.shirt n  
 Intended: ‘expensive inner shirt’

In the literature these items are called bound stems (Dai 1992:40,75–76) or bound roots (Sproat and Shih 1996; Packard 2000; Pirani 2008); see Wang (2018) for a review of the literature. In the present theory, we do not assume a level of stems (Embick and Halle 2005), hence there are no bound stems, and it is potentially misleading to call them bound roots, as roots are, by hypothesis, devoid of categorial features, and are therefore necessarily bound. Instead, following Wang (2018) we call them bound words. At this point two questions call for an answer: (a) Why can these items not occur as free words, and (b) why can nominal bound words not be reduplicated in TXC?

What makes them special, we claim, following Wang (2018), is that they are not made up of a root and a categorizer, but are monomorphemic lexical items with inherent category. In this respect they are like light verbs such as English *go*, *do*, *take*, etc. These are functional categories, yet have some semantic features including assigning theta-role. As functional categories they are monomorphemic, not made up of a

<sup>21</sup>This is a so-called parallel compound, to be discussed below in Sect. 6.2.2.

root and a categorizer. Embick (2015) represents them as *go<sub>v</sub>*, *do<sub>v</sub>* etc. There are no reasons to think that the Chinese bound words are functional light nouns, verbs, and adjectives, though. Light verbs have highly general meanings: causation, initiation, change of state, etc. (see Butt 2010), while the Chinese bound words mostly have quite specific meaning, for example ‘tongue’, ‘film’, ‘duvet’, ‘inner shirt’, to take some nominal examples from the list of words in (37). We maintain, therefore, that they are based on roots, but like light verbs they are monomorphemic: a root with inherent category. Now if noun reduplication in TXC is the copying of the phonological matrix of a root by a sister nominalizer, it follows that bound nouns do not undergo reduplication.

Why can they not form free words? Note first how the absence of reduplication of bound nouns in TXC provides an argument that the boundness condition is not phonological by nature, as claimed in some of the literature; see Chung et al. (2014:613). It is not for instance the case that the items are subject to a phonological condition that they must be part of a disyllabic word. If that were the case, reduplication would seem to be the perfect mechanism to satisfy it in TXC, but reduplication is not even possible.

Following Wang (2018:68), we propose that there is a condition (41) in Chinese, which may even be universal.

- (41) A free content word (= root-based word) must contain at least two morphemes.

‘Content word’ is meant to exclude light verbs and comparable functional heads. In the common case of an acategorial root merged with a categorizer the result is of course a bimorphemic word. The hypothesis put forward by Wang (2018) is that the Chinese bound word is monomorphemic, a root with a categorial feature forming a single item. The principle (41) is, then, the reason why it cannot function as a free word. But once a bound word is merged with another morpheme (an affix, a free word or another bound word) the result is a licit free word, satisfying (41). The one functional category they cannot merge with, though, is a categorizer, since they already have syntactic category.

As mentioned at the start of Sect. 3.2, a categorizer can only merge with an acategorial item. We proposed that this is an effect of Chomsky’s (2013) labelling algorithm: the grammar cannot assign a label to a set made up of a word (= a root with a category) and a categorizer. This applies equally whether the word is itself made up of a root merged with a categorial feature or is a monomorphemic item with a categorial feature, as bound words are, by hypothesis.

We now have answers to the two questions we posed: Why do bound words have to be bound, and why can bound nouns not be reduplicated in TXC? The answer to both questions presupposes that bound words are not made up of a root merged with a categorizer. Conversely, we now have evidence from two distinct phenomena of the analysis of the bound word as a monomorphemic item, a root with inherent category: (a) they cannot form a free word on their own, and (b) they cannot be reduplicated.

Matters are complicated by the observation that bound words can be either the head or the non-head of a compound. Recall that we have claimed, in part based on a comparison with compounding in Swedish, that the non-head of a simple compound



adjective-noun compounds and verb-object compounds such as in (43a,b) (retaining the transcription used in Hu and Perry 2017).

- (43) a. *fao* *ɲɲ* (Yixing Chinese)  
           small person  
           ‘kid’  
       b. *ɲe* *ɲy*  
           read book  
           ‘study’  
           (Hu and Perry 2017:711)

Following Marantz (2007), Embick and Noyer (2007) and Borer (2013), Hu and Perry (2017:718) propose that the NCCs in Yixing Chinese are the result of merging two bare roots which are devoid of syntactic features, followed by merge of a categorizer. Since the bare roots have no syntactic features, the object which is the result of merging two roots has no label. Hu and Perry use the zero symbol  $\emptyset$  for this unlabelled phrase (Hu and Perry 2017:724). The categorization comes after the merge of two roots (Hu and Perry 2017:719), by merging a categorizer with the unlabelled phrase. So for example the NCC *fao* *ɲy* ‘small book: comic’ which is a noun, will have the following structure where two roots make a unit which is merged with a null nominalizer:

- (44)
- 
- ```

      N
     / \
    Ø   n
   /  \
  R    R
 /    \
fao    ɲy
small book
  
```

According to this analysis, although the NCC has a word category, the lexical components of the NCCs do not have inherent word category. The components may thus occur in NCCs with a different category (Hu and Perry 2017:719). The following is an example. The compound here is a so-called synonymy-compound: two synonymous roots form a compound with the same meaning as its components (this subtype of NCCs, will be discussed below in Sect. 6.2.2). In (45a), the unit made up of two merged roots is merged with an adjectival categorizer, thus forming a compound adjective. In (45b) the same two roots are merged with a nominal categorizer, forming a nominal compound.

- (45) a. *gə* *gə* *wodou* *ma* *p^hjaol* *jaŋ* (Yixing Chinese)
 this CLF girl very beautiful beautiful
 ‘this girl is very beautiful’
 b. *p^hjaol* *jaŋ* *dae* *ɲe* *ɲy* *mmə* *joŋ*
 beautiful beautiful read book NEG.have use
 ‘Beauty is of no use in study.’
 (based on Hu and Perry 2017:718)

Hu and Perry's observations and analyses of NCCs are very similar to those of Zhang (2007). Zhang discusses some exocentric compounds in Mandarin which do not have compositional meanings. In these compounds, the category of the components does not match that of the resultant compound, which she branded as an abnormality. For example, she observes that the verb-object compound, a type of compound in Mandarin, can be a noun or an adjective:

- (46) zhī jǐ (Mandarin)
 know self
 'close friend'
 (Zhang 2007:173)
- (47) chī lì (Mandarin)
 eat energy
 'laborious'

According to Zhang (2007), the above abnormality concerning verb-object compounds can be explained if the components of such compounds are roots without syntactic features, while the category of the compound is determined by a third element, which is a category-defining functional head. In this way there will be no discrepancy between the category of the root components (as they have no category) and the category of the compound. So for example (46), which is a noun, will have the following structure where two roots are merged first, then the resultant complex Root + Root will merge with the category-defining functional head *f*. In (46) *f* is nominal, in (47) it is adjectival.

- (48)
-
- (based on Zhang 2007:173)

By standard assumptions, which we have adopted, labelling of syntactic units is needed for linearization: The linear order of a phrase as well as a complex word depends on which constituent is the head, with variation across languages and also within a language depending on the category of the phrase or word. This also determines the semantic interpretation of the phrase or complex word. Also by standard assumptions, widely although not universally accepted, labelling is needed to make a word or phrase visible for further (morpho-) syntactic operations.²³

Another, less well-established, hypothesis is that labelling is needed in order to link a syntactic unit to an encyclopedic entry (Zhang 2007; Bauke 2014:65–70; Hu and Perry 2017). This would be why a compound consisting of two roots is not interpreted compositionally. The bare roots cannot be linked with any encyclopedic entry, but the compound can be as it has a label provided by the categorizer. The meaning of the compound thereby need not have any direct, systematic relation to the meaning that the roots would have if they were labelled. This is an idea which we cannot

²³See Collins (2002) for a rejection of this idea. See Hornstein (2009:46) for some discussion.

At the outset of Sect. 3 we stated that the categorizer *n* can only merge with a labelless (= acategorial) construct, typically with a root. Two merged roots is still a labelless construct, and hence can be merged with *n*. We will represent the root-root compound as labelless.²⁵

6.2.1 Non-compositional ‘attributive’ compounds

(49) a. jiéng fóng (TXC)
quiet room
'place for practising Buddhism'

b. chǒng tuēi
long leg
'person who does not mind long walks'

²⁵ This is what Hu and Perry's (2017) 'Ø' in (44) represents, as well. We find Zhang's (2007) label R+R in (48) redundant, as it replicates what the tree already represents: two merged roots.

The above compounds have idiosyncratic meanings and are nouns. As NCCs, (49a) and (49b) do not have reduplicated forms. (49a) cannot be reduplicated as (50a), and (49b) cannot be reduplicated as (50b). The complete compound cannot be reduplicated, either, as shown in (50c) and (50d).

- (50) a. *jiěng fǒng fǒng (TXC)
 quiet room n
 b. *chǒng tuēi tuēi
 long leg n
 c. *jiéng fǒng jiéng fǒng
 d. *chǒng tuēi chǒng tuēi

Recall that the presence of the categorizer n is a precondition for reduplication to occur in TXC. So apparently the object-referring items *fǒng* ‘room’ and *tuēi* ‘leg’ in (49) are not words which contain a categorizer. But as free words they can be, and need to be, reduplicated: *fǒng fǒng* ‘room’, *tuēi tuēi* ‘leg’. And as we would also predict, as heads of compositional compounds, they can be reduplicated:

- (51) a. fēi fǒng fǒng (TXC)
 water room n
 ‘water room’
 b. zhū tuēi tuēi
 table leg n
 ‘table leg: leg of table’

So *fǒng* ‘room’ and *tuēi* ‘leg’ in (49) are roots. *Jiéng* ‘quiet’ and *chǒng* ‘long’ are roots which may merge with a categorizer to form words, but in (49) they can only be roots. We know this, because as words they would be heads in (49), which they are not.

(49a) and (49b) are nouns, so they will have the following structure:

If an object-referring item is the head in a compound in TXC and is not a bound word, then it can undergo reduplication. If it is not the head, it is a root and cannot undergo reduplication, as we showed in Sect. 4.2. The absence of reduplication seen in (49) is therefore predicted if the structure of the NCCs is (52): The categorizer cannot copy the phonological matrix of either of the roots, as they are not sisters. As for why it cannot copy the phonological matrix of the unit made up of the two roots, we proposed in Sect. 3.2 a condition on reduplication prohibiting reduplication of more than one syllable. This rules out reduplicating the root-root unit in (52) to derive (50c,d).

The structure which is proposed here for NCCs is in line with what Hu and Perry (2017) and Zhang (2007) have proposed. In TXC there is now additional direct empirical evidence, that is absence of reduplication, to support the analysis.

6.2.2 Non-compositional coordinative compounds

Synonymy-compound nouns and parallel compound nouns, which are two subtypes of coordinative compounds in Mandarin and TXC, are introduced and analysed in this section (see Wang 2018 for discussion of other types of non-attributive compounds in Mandarin and TXC). Synonymy-compounds are compounds in which the components are synonyms (Chao 1968:374–375; Ceccagno and Basciano 2011:481). Examples in Mandarin and TXC are demonstrated here:

- (53) a. hào mǎ (Mandarin)
 number number
 ‘number’
 b. guǎng bó (TXC)
 stick stick
 ‘stick’

Parallel compounds are a type of coordinative compounds where the components are semantically associated, but are neither antonyms nor synonyms (Chao 1968:377). The following examples are parallel compound nouns in Mandarin and TXC:

- (54) a. shuǐ guǒ (Mandarin)
 water fruit
 ‘fruit’
 b. bǎn jǐng (TXC)
 board neck
 ‘neck’

We have argued that reduplication applies where a root is merged with a nominalizer. This implies that a constituent in a compound which cannot be reduplicated will not have that structure. As a matter of fact, the constituents of synonymy and parallel compounds cannot be reduplicated. We take this to be evidence that the constituents of these compounds, when they are not bound words, are roots, not nouns. This also tallies with the fact that the compounds are non-compositional. This is clearer in the case of parallel compounds, and we therefore begin with a discussion of parallel compound nouns. The following is a list of parallel compound nouns in TXC.

- (55) a. niān wù (TXC)
 eye pit
 ‘eye’
 b. bǎn jǐng (TXC)
 board neck
 ‘neck’

The meaning of these compounds is clearly not the sum of their parts, and there is no systematic semantic relation between the two parts. For instance, *niān* ‘eye’ in (55a) does not modify *wù* ‘pit’, or bear any other definable semantic relation to it; the same is true for *bǎn* ‘board’ and *jǐng* ‘neck’ in (55b). The meaning and linear order of the compounds is idiosyncratic, represented as such in the Encyclope-

dia. Following Zhang (2007), Hu and Perry (2017), and Wang (2018), this is explained if components of parallel compounds in TXC are two merged roots, which are subsequently merged with a nominal categorizer, at which point it can be linked with an entry in the Encyclopedia. The structure of, for example, (55a) would be (56).

The components of (55a) cannot be reduplicated, nor can the entire compound be reduplicated:

- (57) a. **niān niān wū* (TXC)
 eye n pit
 b. **niān wū wū*

As independent words, *niān* ‘eye’ and *wū* ‘pit’ can, and must, undergo reduplication, so in this case they are roots merged with a null nominalizer: *niān niān* ‘eye’, *wū wū* ‘pit’. We conclude from this that the components of the parallel compound (55a) are bare roots. *Bān* ‘board’ in (55b) is a root, as well, as it cannot be reduplicated, while it can be and must be reduplicated as a free word. As for *jiěng* ‘neck’ in (55b), it is a bound word. As such it cannot be reduplicated in any context, including (55b), and cannot occur alone as a free word in phrases. As discussed in Sect. 5, bound words occur in two guises, either as roots with inherent category or as bare roots. In the parallel compounds it must be the bare root version or they would label the compound, which would thus be compositional, contrary to fact. Thus (55b) will have the following structure:²⁶

Next, consider the synonymy compound (53b). As shown in (59), neither constituent can be reduplicated.

²⁶Packard (2000:195) assumes that parallel compounds have a structural head, namely the constituent on the right, which is sometimes different from the semantic head. Reduplication in TXC provides evidence against this analysis.

- (59) a. *guěng guēng bóng (TXC)
 stick n stick
 b. *guéng bǒng bōng
 stick stick n

So neither *guéng* ‘stick’ nor *bóng* ‘stick’ in (53b) is a noun made up of a root and a null nominalizer. As independent words *guéng* ‘stick’ and *bóng* ‘stick’ can be, and must be, reduplicated: *guěng guēng*, *bǒng bōng*. This means that they are able to merge with the null nominalizer. The reason why they cannot be reduplicated in the synonymy compound (53b) is thus, most plausibly, that they are bare roots, not merged with a null nominalizer. (53b) would have the following structure, the standard NCC structure.

- (60)
-

Similar to parallel compounds in TXC, synonymy compounds in TXC can also be combinations of bound words and roots.

- (61) tǔ mō (TXC)
 saliva saliva
 ‘saliva’
- (62) bó géng
 neck neck
 ‘neck’

Tǔ ‘saliva’ cannot occur alone as a free word in a phrase, nor can it ever be reduplicated alone, so it is a bound word. Given the fixed word order of (61), it must be the root version of the bound word, or it would be head, but in that case it would be the rightmost item. We assume the same is true of *géng* ‘neck’ in (62), although in this case it is not as easy to show that it is not the head of the compound. As for *mō* ‘saliva’ and *bó* ‘neck’, they are roots, based on the fact that they cannot be reduplicated in (61) and (62) respectively, but can be reduplicated alone. So the structure of the synonymy compounds (61) and (62) would be the same as for parallel compounds and other NCCs, that is (60).

We assume that synonymy-compound nouns and parallel compound nouns are formed by merge of roots or bound words in Mandarin, as well, even though we do not have the direct evidence of it that we see in TXC. For instance, *hào* ‘number’ and *mǎ* ‘number’ in (53a), *shuǐ* ‘water’ in (54a) can each occur as a free word, which is to say, they are roots that can merge with a null nominalizer to form a free word. However, we propose, by analogy with the corresponding compounds in TXC, that the compounds in (53a) and (54a) are formed by merge of two bare roots, followed by merge of a null nominal categorizer.

The order of the constituents of the parallel and synonymy compounds is fixed. The opposite order of the constituents in (53a,b) and (54a,b), for example, yields ungrammatical compounds. In compositional compounds the order is determined by

headedness (see Sect. 3.1); in Chinese, the head is spelled out rightmost. In NCCs, being headless, the order as well as the interpretation is not derived but encoded in the Encyclopedia.

7 Recursion in word formation

In Sect. 3 we argued that an attributive compound with two members could not consist of two merged words, for instance two merged nouns, as the grammar could not then determine which one is the head, i.e. which one will label the set. A solution is to form a compound by merging a root with a noun, in which case only the noun can be head. This is evident in a class of nouns in Swedish which overtly consist of a root plus a vowel serving as nominalizer. We have shown that it is also evident in the case of attributive compounds generally in TXC, as in this language nouns, as opposed to roots, are always reduplicated by virtue of the nominalizer copying the phonological matrix of the root. However, compounds can be more complex. In compounds such as *wallpaper design* and *football pitch* the modifier is itself a compositional compound, hence not a root but a noun. So these compounds do consist of two nouns, one of which is itself a compound. They exemplify recursion, more precisely left-recursion, in compounds. How can the grammar determine which noun is the head?

As discussed in Holmberg (1992) and Josefsson (1997, 1998), in Swedish this is fixed by assigning Case to the complex modifier, mostly realized as a suffix -s, historically related to the genitive Case used in possessive NPs. Compare the simple compounds in (63a) with the recursive compounds in (63b).

- (63) a. boll-plan, flick-väska, (Swedish)
ball-pitch girl-bag
b. fot -boll-s -plan, skol -flick-s -väska
foot ball-CASE-pitch school-girl-CASE-bag ('a bag for schoolgirls')

The structure of for example *skolflicksväska* would be (64).

According to Holmberg (1992), the effect of the Case is to close the nominal projection, preventing projection of the features of the Case-marked noun onto the dominating node, thus allowing the other noun to project its features, and label the set made up of the compound and the simple noun. The generalization in Swedish is that the modifier of an attributive compound is either a root or it is Case-marked.

TXC has a different way of solving the problem: It has no recursive compounding. The compounds in (65), for example, where (65a) has a compositional compound and (65b) an NCC as non-head, are ungrammatical.

- (65) a. **[liěng hǔ] sēi* (TXC)
 collar rim colour
 Intended reading: ‘colour of rim of collar’
 b. **[niǎn wū] sēi*
 eye pit colour
 Intended reading: ‘eye colour’

There are languages that have endocentric attributive compounds but where recursion is either rare or non-existent. Greek is one example (Ralli 2013:95), Polish is another Szymanek (2017).²⁷ We can add TXC to the list. We conclude that TXC, along with at least some of the other languages which lack recursion, only allows roots as non-heads of compounds.

Mandarin, along with many other languages including English, allows recursion and shows no overt evidence of any special marking of non-heads of compounds. How to accommodate this with the present theory is something we leave for future research (see Mukai 2008 for a proposal). One possibility, suggested in Holmberg (1992), is that there are two different ways to form compounds allowed by UG. One is by means of set merge in Chomsky’s (2001) sense, as outlined in Sect. 3.1. The other is by ‘pair merge’ (Chomsky 2001), that is adjunction, an inherently asymmetric process, where labelling is therefore not a problem: adjunction of α to β does not create a new category, only expands β . Holmberg (1992) suggests that Swedish is of the former type, English of the latter type, supported by the fact that English allows inflection of the non-head, as in *arms race* and *parks attendant*, which is never allowed in Swedish. Mandarin would be of the pair merge type, allowing recursive compounding without any special marking of complex non-heads, while TXC would be of the set merge type along with Swedish, but as it lacks means to case-mark complex non-heads the effect in TXC would be the absence of recursive compounding.

8 Conclusions

In Xining Chinese, particularly as spoken by the elderly, what we have called Traditional Xining Chinese (TXC), free nouns are always reduplicated, as a purely formal requirement with no effect on interpretation. The relevant descriptive generalization can be stated as follows:

²⁷ Anders Holmberg and Samuel Andersson have collected data on compounding from 208 languages on the basis of descriptive grammars. Typically there is no mention of whether compounding is recursive or not. In fact, only in *one* grammar was absence of recursive (nominal) compounds explicitly mentioned: Hellenthal (2010: 180) on Sheko, an Omotic language spoken in Ethiopia.

There are languages that have compound words, but only exocentric ones, including the Romance languages and Arabic; cf. French *essui-glace* lit. ‘wipe-glass,’ i.e. ‘windscreen wiper.’ These are not composed of a root and a head, in the manner of nominal compounds in Chinese, English, and other languages discussed in the present article. They may be constructed as NCCs in Chinese and other languages. The compounds in Sheko appear to be of the endocentric type.

- (66) a. Reduplicate a free, simple noun.
 b. Optionally reduplicate the head of a complex noun, except if it is functional or a bound word.

This is the effect of an operation, a language-particular property of TXC, which copies the phonological features of a monosyllabic root onto a sister categorizer *n*. We have referred to it as noun reduplication. The reduplication is optional, but forced in the case of free monosyllabic nouns by an output condition or filter:

- (67) *N if N is a free word and has less than two syllables

There is a condition restricting the reduplication which is likewise sensitive to syllable structure: It can only apply to a syllable. This rules out reduplication of root-root compounds (which are by definition non-compositional, in the present theory) and polysyllabic roots.

Given this, every part of the generalization in (66) is explained by the theory we articulate, building on Chomsky (1995, 2013), much work in the Distributive Morphology tradition, and recent work on the morphology of Mandarin and TXC in Wang (2018). An important theoretical premise is that roots are acategorical. Content words are derived by merging a root with a categorizer, which may be, and often is, spelled out as null (Marantz 1997; Josefsson 1997, 1998; Harley and Noyer 1999; Lieber 2006; Embick and Noyer 2007, 2008; Harley 2011, 2014; de Belder 2011; Hu and Perry 2017). The reduplication in TXC is the result when a nominal categorizer copies the phonological matrix of its sister root, thus meeting the condition that a noun must have two syllables/two pronounced morphemes.

The reason why the non-head of a compound or derived word cannot be reduplicated is that it is a root, merged with a word (compounds) or with an affix other than a categorizer (derived words). A root without a categorizer as sister cannot reduplicate. Functional heads cannot reduplicate, as they are single items with inherent category. As we show, there are also acategorical affixes, though, which merge with nouns (not roots). These nouns therefore can undergo reduplication in TXC but need not to, as they satisfy (67) even without reduplication.

Bound words is a class of words typical of Chinese including TXC. Like function words they do not consist of a root with a categorizer, but are single items with inherent category (Wang 2018). As such they cannot undergo reduplication. Due to the output condition (41), they have to be morphologically bound, though, either by affixation or by compounding.

In NCCs there is no reduplication. Following work by Zhang (2007), Bauke (2014), and Hu and Perry (2017) we assume that NCCs are made up of two merged roots, forming an unlabelled unit which is merged with a categorizer. This configuration does not meet the structural conditions for reduplication in TXC, as the null categorizer is not a sister of either root, and the unit consisting of two roots cannot be reduplicated, as reduplication can only apply to a single syllable.

We have assumed (following Holmberg 1992 and Josefsson 1997, 1998) that simple compound nouns typically consist of a root and a noun, the latter the head. This conforms to the principle that units built by merge must be asymmetric for labelling to be possible (Chomsky 2013), and explains why the non-head of a simple com-

pound cannot be reduplicated in TXC. This raises the question of how complex recursive compounds are derived, where the non-head is itself a compound. The issue is avoided in TXC, as the language does not allow recursive compounding.

A question we have not touched upon yet is why only nouns undergo semantically vacuous reduplication in TXC. Nothing in the theory we have articulated predicts this. It is suggestive, though, that, as discussed by Duanmu (1999), most nouns in the common vocabulary of Mandarin are disyllabic or more; only 16% are monosyllabic. The percentage of monosyllables is clearly higher for verbs (41%) and for adjectives (31%). Whatever the explanation is of this distribution, the restriction of reduplication to nouns in TXC is an extreme version of the same pattern: In TXC no nouns are monosyllabic.

As mentioned, the obligatory noun reduplication we have investigated is characteristic of the variety of Xining Chinese spoken by the older generation, Traditional Xining Chinese (TXC). For younger speakers, presumably as an effect of the increasing influence of Standard Mandarin, reduplication is common but not compulsory. Just what the factors are which govern the variation between reduplicated and non-reduplicated forms—grammatical, lexical, and social factors—is an interesting topic for future research, which may itself shed more light on the formal nature of the reduplication.

Considering the place of the reduplication discussed here in the architecture of the grammar, it is interesting to compare with Travis's (2001) findings. She describes three cases of reduplication. None of them is a close relative of the reduplication we have described here. In particular, they all have semantic effects (on quantity, one way or other), and they all apply in the phrasal domain, across a maximal category boundary. Her conclusion is still interesting and relevant for the reduplication described here: "... all cases of reduplication [discussed in her paper] are filling (by creating copies) positions that are independently available in the syntax." (Travis 2001:11) This makes them different from the cases of reduplication described and discussed in the (morpho-) phonological literature (see Frampton 2009), which may be triggered by a null functional head, but which perform a phonological operation typically on an adjacent vocabulary item realizing a root, expanding that item by reduplicating it, or more often, a part of it (see Frampton 2009). In this perspective, noun reduplication in Xining Chinese is more like Travis's (2001) cases of reduplication, more 'syntactic' in the way it employs copying of the phonological matrix of a vocabulary item onto an independently generated head.

Acknowledgements Thanks to the organisers and audiences at CamCoS 7, Olinco 2018, and ISOCTAL 3 at University College Cork. Special thanks to Rint Sybesma, Alec Marantz, and S.J. Hannahs for comments, critique and advice. Thanks also to the anonymous reviewers of *NLLT* and especially the editor Daniel Harbour.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material.

If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Basciano, Bianca, and Antonella Ceccagno. 2009. The Chinese language and some notions from Western linguistics. *Lingue et Linguaggio* VIII.1: 105–135.
- Bauke, Leah. S. 2014. *Symmetry breaking in syntax and the lexicon*. Amsterdam: John Benjamins Publishing Company.
- Bell, Daniel. 2017. Syntactic change in Xining Mandarin. PhD diss., Newcastle University.
- Bell, Daniel. 2019. Chinese possessors Japanese style scrambling: The case of the Xining dialect. *Journal of East Asian Linguistics* 28: 143–173.
- Bobaljik, Jonathan. 2000. The ins and outs of contextual allomorphy. *University of Maryland Working Papers in Linguistics* 10: 35–71.
- Borer, Hagit. 2005. *Structuring sense, vol. 1: In name only*. Oxford: Oxford University Press.
- Borer, Hagit. 2013. *Structuring sense, vol. 3: Taking form*. Oxford: Oxford University Press.
- Borer, Hagit. 2014a. The category of roots. In *The syntax of roots and the roots of syntax*, eds. Artemis Alexiadou, Hagit Borer and Florian Schäfer, 112–148. Oxford: Oxford University Press.
- Borer, Hagit. 2014b. Wherefore roots? *Theoretical Linguistics* 40: 343–359.
- Butt, Miriam. 2010. The light verb jungle: Still hacking away. In *Complex predicates: Cross-linguistic perspectives on event structure*, eds. Mengistu Amberber, Brett Baker and Mark Harvey, 48–78. Cambridge: Cambridge University Press.
- Ceccagno, Antonella, and Bianca Basciano. 2011. Sino-Tibetan: Mandarin Chinese. In *The Oxford handbook of compounding*, eds. Rochelle Lieber and Pavol Štekauer, 477–490. Oxford: Oxford University Press.
- Chao, Yuanren. 1968. *The grammar of spoken Chinese*. Berkeley: University of California Press.
- Chomsky, Noam. 1965. *Aspects of a theory of syntax*. Cambridge: MIT Press.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: a life in language*, ed. Michael Kenstowicz, 1–53. Cambridge: MIT Press.
- Chomsky, N. 2008. On phases. In *Foundational issues in linguistic theory*, eds. Robert Freidin, Carlos Otero, and Maria-Luisa Zubizarreta, 133–166. Cambridge: MIT Press.
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130: 33–49.
- Chung, Karen S., Nathan Hill, and Jackson T-S. Sun. 2014. Sino-Tibetan. In *Oxford handbook of derivational morphology*, eds. Rochelle Lieber, and Pavol Štekauer, 609–634. Oxford: Oxford University Press.
- Collins, Christopher. 2002. Eliminating labels. In *Derivation and explanation in the minimalist program*, eds. Samuel D. Epstein and T. Daniel Seely, 42–64. Cambridge: Cambridge University Press.
- Creemers, Ava, Jan Don, and Paula Fenger. 2018. Some affixes are roots, others are heads. *Natural Language and Linguistic Theory* 36: 45–84.
- Dai, Xiangling. 1992. Chinese morphology and its interface with the syntax. PhD diss., Ohio State University.
- Dede, Keith. 1993. *Language contact on the Qinghai-Gansu border area*, Masters thesis, University of Washington.
- Dede, Keith. 1999. Language contact, variation and change: The locative in Xining, Qinghai. PhD diss., University of Washington.
- Dede, Keith. 2003. The Chinese language in Qinghai. *Studia Orientalia* 95: 321–346.
- Dede, Keith. 2006. Standard Chinese and the Xining dialect: The rise of an interdialectal standard. *Journal of Asian Pacific Communication* 16: 319–334.
- De Belder, Marijke. 2011. Roots and affixes: Eliminating lexical categories from syntax. PhD diss., Utrecht University.
- Di Sciullo Anna Maria. 2005. *Asymmetric morphology*. Cambridge: MIT Press.
- Duanmu, San. 1999. Stress and the development of disyllabic vocabulary in Chinese. *Diachronica* 16: 1–35.

- Ejerhed, Eva, and Hank Bromley. 1986. A self-extending lexicon: Description of a word-learning program. In *Papers from the fifth nordic conference of computational linguistics*, ed. Fred Karlsson, 59–72. Helsinki: University of Helsinki. Publication no. 15.
- Embick, David. 2015. *The morpheme: A theoretical introduction*. Boston: De Gruyter Mouton.
- Embick, David, and Morris Halle. 2005. On the status of stems in morphological theory. In *Romance languages and linguistic theory 2003: Selected papers from 'Going Romance'*. 2003, Nijmegen, 20–22 November, eds. Twan Geerts, Ivo van Ginneken and Haike Jacobs. 37–62. Amsterdam: John Benjamins.
- Embick, David, and Rolf Noyer. 2007. Distributed morphology and the syntax/morphology interface. In *The Oxford handbook of linguistic interfaces*, eds. Gillian Ramchand and Charles Reiss, 289–324. Oxford: Oxford University Press.
- Embick, David, and Rolf Noyer. 2008. Architecture and blocking. *Linguistic Inquiry* 39: 1–53.
- Frampton, John. 2009. *Distributed reduplication*. Cambridge: MIT Press.
- Hale, Ken, and Jay Keyser. 2002. *Prolegomena to a theory of argument structure*. Cambridge: MIT Press.
- Hall, T. Alan. 1999. The phonological word: A review. In *Studies on the phonological word*, eds. Tracy A. Hall and Ursula Kleinhenz, 1–22. Amsterdam: Benjamins.
- Halle, Morris, and Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The view from Building 20*, eds. Ken Hale and Jay Keyser, 111–176. Cambridge: MIT Press.
- Harley, Heidi. 2004. Merge, conflation and head movement: the First Sister Principle revisited. In *Proceedings of NELS 34*, eds. Kier Mouton and Matthew Wolf, 239–254. GSLA: U. Mass Amherst.
- Harley, Heidi. 2011. Compounding in distributed morphology. In *The Oxford handbook of compounding*, eds. Rochelle Lieber and Pavol Štekauer, 129–144. Oxford: Oxford University Press.
- Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40: 225–276.
- Harley, Heidi, and Rolf Noyer. 1999. Distributed morphology. *Glott International* 4: 3–9.
- Harris, James W. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22: 27–62.
- Hellenthal, Anne-Christie. 2010. A grammar of Sheko. Ph.D. diss., University of Leiden. Published by LOT, Utrecht.
- Höhn, George. 2017. *Non-possessive person in the nominal domain*, PhD dissertation, University of Cambridge.
- Holmberg, Anders. 1992. Properties of non-heads in compounds: A case study. *Working paper in Scandinavian syntax* 49: 27–58.
- Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and universal grammar*. Cambridge: Cambridge University Press.
- Huang, Churen, and Dingxu Shi. 2016. Preliminaries. In *A reference grammar of Chinese*, eds. Chu-Ren Huang and Dingxu Shi, 1–14. Cambridge: Cambridge University Press.
- Huang C.T. James, Y.-H. Audrey Li, and Yafei Li. 2009. *The syntax of Chinese*. Cambridge: Cambridge University Press.
- Hu, Xuhui, and J. Joseph Perry. 2017. The syntax and phonology of non-compositional compounds in Yixing Chinese. *Natural Language and Linguistic Theory* 36: 701–742.
- Josefsson, Gunlög. 1997. On the principles of word formation in Swedish. PhD diss., Lund University.
- Josefsson, Gunlög. 1998. *Minimal words in a minimal syntax*. Amsterdam: Benjamins.
- Kiefer, Ferenc. 1970. *Swedish morphology*. Stockholm: Skriptor.
- Lehmann, Christian. 2010. Roots, stems, and word classes. In *Parts of speech: Empirical and theoretical advances*, eds. Umberto Ansaldi, Jan Don, and Roland Pfau, 43–64. Amsterdam: Benjamins.
- Li, Y.-H. Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics*. 8: 75–99.
- Li Zhenjie and Yukun Bai (eds.) 1987. *Zhongguo baokan xin ciyu [New Chinese press terms]*. Beijing: Huayu Jiaoxue Chubanshe.
- Lieber, Rochelle. 1980. On the organization of the lexicon. PhD diss., MIT
- Lieber, Rochelle. 2006. The category of roots and the roots of categories: What we learn from selection in derivation. *Morphology* 16: 247–272.
- Lowenstamm, Jean. 2015. Derivational affixes as roots: Phrasal spell-out meets English stress-shift. In *The syntax of roots and the roots of syntax*, eds. Artemis Alexiadou, Hagit Borer, and Florian Schäfer, 230–259. Oxford: Oxford University Press.
- McCarthy, John J., and Alan Prince. 1990. Foot and word in prosodic morphology: The Arabic broken plurals. *Natural Language and Linguistic Theory* 8: 209–282.
- Marantz, Alec. 1997. No escape from syntax: Do not try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2): 201–225.
- Marantz, Alec. 2007. Phases and words. In *Phases in the theory of grammar*, ed. Sook-Hee Choe, 191–222. Seoul: Doong-In Publishing Co.

- Moro, Andrea. 2000. *Dynamic antisymmetry*. Cambridge: MIT Press.
- Mukai, Makiko. 2008. Recursive compounds. *Word Structure* 1: 178–198.
- Packard, Jerome L. 2000. *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge: Cambridge University Press.
- Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.
- Pirani, Laura. 2008. Bound roots in Mandarin Chinese and comparison with European semi-words. In *20th North American Conference on Chinese Linguistics (NACCL-20)*, 1:261–277.
- Po-Ching, Yip, and Don Rimmington. 2016. *Chinese: A comprehensive grammar*. New York: Routledge.
- Postal, Paul. 1969. On so-called “pronouns” in English. In *Modern studies in English*, eds. David A. Reibel and Sanford A. Schane, 201–224. Englewood: Prentice Hall.
- Ralli, Angela. 2013. *Compounding in modern Greek*. New York: Springer.
- Ren, BiSheng. 2006. *Qinghai fangyan yufa zhuanji yanjiu [The grammar of the Qinghai dialect]*. Xining: Qinghai Renmin Chubanshe.
- Sproat, Richard, and Chilin Shih. 1996. A corpus-based analysis of Mandarin nominal root compound. *Journal of East Asian Linguistics* 5: 49–71.
- Szymanek, Bogdan. 2017. Compounding in Polish and the absence of phrasal compounding. In *Investigations into the nature of phrasal compounding*, eds. Carola Trips and Jaklin Kornfilt, 47–79. Oxford: Oxford University Press.
- Travis, Lisa. 2001. The syntax of reduplication. In *Northeast Linguistics Society (NELS)*, 31: 454–469. Amherst: GLSA.
- Ueda, Yasuki, and Tomoko Haraguchi. 2008. Plurality in Japanese and Chinese. *Nanzan Linguistics (Special Issue)*, 229–242.
- Wang, Shuangcheng. 2009. Xining fangyan chongdieshi [Reduplication in Xining Chinese]. *Qinghai shifan daxue minzu shifan xueyuan xuebao* 5: 1–4.
- Wang, Qi. 2018. The structure of nouns in Old Xining and Modern Standard Chinese. PhD diss., Newcastle University.
- Wang, Shuangcheng, and Keith Dede. 2016. Negation in the Xining dialect. *Language and Linguistics* 17: 403–429.
- Williams, Edwin. 1981. On the notions lexically related and head of a word. *Linguistic Inquiry* 12: 245–274.
- Zhang, Chengcai. 2001. *Qinghai shengzhi (fangyan zhi) [Qinghai annals (dialect annals)]*. Qinghai: Huangshan.
- Zhang, Chengcai, and Shikuei Zhu. 1987. *Xining fangyanzhi [Xining dialect annal]*. Xining: Qinghai Renmin Chubanshe.
- Zhang Niina Ning. 2007. Root merger in Chinese compounds. *Studia Linguistica* 2: 170–184.
- ZWGW (Zhongguo Wenzi Gaige Weiyuanhui Yanjiu Tuiguang Chu [Chinese Language Reform Committee Research and Popularization Office]). 1959. *Putonghua san qian chang yong ci biao [Three thousand and commonly used words in Standard Chinese]*. Beijing: Wenzi Gaige Chubanshe. Preliminary edition.