Roots, categorizers and reduplication in Xining Chinese

[[a]**Anders Holmberg** and [b]**Qi Wang**]

[Newcastle University, Newcastle upon Tyne, UK]

[[a]anders.holmberg@newcastle.ac.uk; [b]qwssg17@outlook.com]

1. Introduction

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

(1)   (a)

Nao	sa	da	zi	**fo**	**fo**	ha	yo	-go
I	PRT	big	DE	spoon	spoon	OBL	need	-AFF

'I need a big spoon'

    (b)

Jia	sa	**mo**	**mo**	ha	mei	ha	zhei
She	PRT	steamed bun	steamed bun	OBL	buy	PRT	PRT

'She has bought steamed buns'

    (c)

Zhi	go	**hai**	**hai**	guei	zhao	guei	zhao
this	CL	shoe	shoe	expensive	PRT	expensive	PRT

'This pair of shoes are 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 North West China. Our data are exclusively from Traditional Xining Chinese, though, abbreviated TXC (in Wang 2018 the dialect is called Old Xining Chinese, OXC).[1]

The analysis we propose here is that the reduplication of nouns in TXC is the result of 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. The categorizer is often a null morpheme. That is the case with lexical categories generally in for example 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 always overt. If it is not realized as an affix, it is realized by copying the phonological features of its sister root.

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 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 a category-less root merged with a designated categorizer. This hypothesis is widely but not universally assumed within generative morphosyntax (see Borer 2014 for a rebuttal), and is even more controversial in more traditional morphological theory.

An important premise is that the reduplication, although it obviously has a phonological effect, is not a phonological operation in the sense of being motivated by phonological conditions and relying on phonological primitives, but is a morphological/morphosyntactic operation. It is not, for example, motivated by conditions on the size of minimal words in TXC (McCarthy and Prince 1990; Hall 1999). For one thing, it concerns specifically nouns, a morphosyntactic, not phonological, category. Verbs and adjectives are not usually reduplicated in TXC, and if they are, it has a semantic effect, denoting repetition or intensification, among other effects. We will demonstrate that the properties and distribution of the reduplication can be predicted under a morphosyntactic approach, but not under a phonological approach.

---

[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. For other work on Xining Chinese, see Dede (2006), Ren (2006), Bell (2017). For other work on reduplication in Xining Chinese, see Ren (2006) and Wang (2009).

[2] The point that reduplication can serve as a probe into the structure of words and phrases is also made by Travis (2001). It could be noted that the reduplication that we describe does not fall into any of the classes of reduplication that Travis identifies.

2. Roots and Categorizers: the Structure of Nouns

2.1. Merge, Labelling, and the Structure and Linear Form of Words

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:

(2) Merge α and β to form a set {α, β} 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):

(3) a.

α
/ \
α   β

b.

β
/ \
α   β
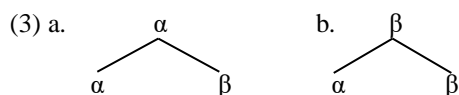
That α and β make up a set, rather than a pair, means that they are not linearly ordered by Merge. Linearization is determined by a phonological rule taking a labelled set as input, so labelling of the set formed by Merge is crucial not only for its interpretation but also its linear order. The rule that is followed in TXC is the same as in English and Mandarin, a version of the Righthand Head Rule of Williams (1980):

(4) A set {α, β} where α is the head projecting a word is linearized as β>α.

Following much work in generative morphosyntax, we assume that common nouns are made up of a root merged with a nominalizer (Josefsson 1997, 1998, Marantz 1997, Harley and Noyer 1999; Embick and Noyer 2007, 2008, Harley 2011, de Belder 2011, Hu and Perry 2017). For a set made up of a root and a nominalizer, the nominalizer will invariably be the head, because the root, by hypothesis, has no categorial or other syntactic features, and thus cannot label the set. It follows that nominalizing affixes in TXC, Mandarin, and English are suffixes.

In derived words, such as, in English, *likeable*, *greatness*, *obesity*, etc., the suffix *-able*, *-ness*, *-ity* is the head, determining the category of the word. The linear order follows from (4). Their status as heads follows directly if the other member of the set is a root.[3]

What about compounds? In a compound such as *wallpaper*, *paper* is the head, determining the interpretation of the compound as denoting a kind of paper, while (4) determines the linear order *wall>paper*. We propose that there are essentially two ways that a set {α, β} making up a compound can be labelled: One of the members, say α, is a word, hence has syntactic category, and β is either a root, in which case it cannot be head, or is a word marked as non-head. Overt marking of a non-head is seen in English compounds such as *men's room* and *bird's nest*. We assume, following Mukai (2008, 2017), that the marking can be, and often is, covert, cross-linguistically. Overt indication that the non-head member of a compound is a root is seen in Swedish compounds such as in (5):

(5) skol-   flicka,   skol-   väska,   flick-   skola,   väsk-   ryckare,   (Swedish)
     school girl   school bag   girl   school   bag   snatcher

The nouns *skola* 'school', *flicka* 'girl, *väska* 'bag', all belonging to the so called 1st declension, are made up of a root (*skol-, flick-, väsk-*) and an overt nominalizer *–a* (also encoding singular number) (Kiefer 1970, Holmberg 1992 and Josefsson 1997, 1998). The non-head of the compounds is the root, while the head is a word itself made up of a root and a nominalizer (*skola, flicka, väska*). (4) determines the linear order as root>word.

As predicted, the root form of these nouns also shows up in derived words, such as *skol-ning* 'schooling' and *flick-aktig* 'girlish', here merged with a derivational suffix functioning as head, determining the syntactic category of the resulting word.

Concluding, the structure of a simple common noun in English is (6a), the structure of a Mandarin common noun in (6b), and the structure of a Swedish common noun of the 1st declension is (6c).

(6) a.

N
/ \
R   n
| |
school   Ø

b.

N
/ \
R   n
| |
shao   Ø

c.

N
/ \
R   n
| |
skol   a

---
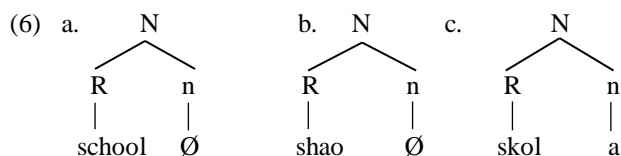
[3] In a framework which in other respects is similar to ours, Josefsson (1998) argues that derivational affixes are themselves roots. This is also argued, and debated, in more recent work: de Belder (2011), Lowenstamm (2015), Creemers, Don, and Felser (2018). We acknowledge that there may be cases where assumed suffixes are actually words, but we reject this analysis for derivational suffixes in TXC, as will be discussed in the text.

2.2. The Structure of Common Nouns in Traditional Xining Chinese

Free common nouns in TXC are always reduplicated; see (1). As in other languages, a common noun in TXC is made up of a root and a nominalizer. The nominalizer is initially null. However, we propose that the null nominalizer in TXC has the characteristic property of copying the phonological features of the sister root, deriving a reduplicated noun. Thus, before reduplication, the structure of the noun 'spoon' is (7a), and after, it is (7b).

(7) a.   N   b.   N
          ╱ ╲   ╱ ╲
         R   n   R   n
         | | | |
         fo   Ø   fo   fo
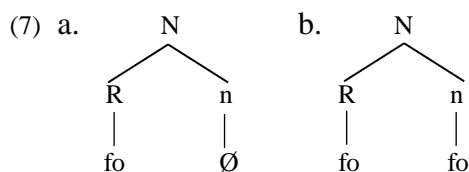
To be more precise, we propose that there is a condition on word structure which may be universal, or else holds for a class of languages including the languages mentioned so far: English, Swedish, Mandarin and TXC, which is (8); we will refer to it as the two-constituent condition

(8)   A content word is made up of minimally two constituents.

A special case of a minimal word is content words consisting of a root and a categorizer. As we shall see in section 4, there are content words which do not consist of a root and a categorizer, but satisfy the two-constituent condition in other ways.

TXC has a special version of the two-constituent condition applying to nouns.

(9)   TXC: A noun is made up of at least two pronounced constituents.

This condition is what motivates the reduplication in nouns which do not satisfy the two-constituent condition in other ways. A null nominalizer is ruled out as it would lead to a violation of (9).

The two-constituent condition applies to content words only. There is little reason to think that function words (complementizers, tense and aspect particles, articles, classifiers, etc.) consist of two constituents. As we will demonstrate in section 3.1, for instance nominal suffixes do not undergo reduplication in TXC. Proper names also do not consist of a root and a nominalizer, and are not generally reduplicated (although they can be, especially as pet names).[4] The structure of pronouns is a controversial issue (cf. Cardinaletti and Starke 2000, Dechaine and Wiltschko 2002), which we will not discuss here, except to note that the fact that they cannot be reduplicated in TXC follows if they are not made up by a root and a nominalizer.

3. Predictions

We have proposed that reduplicated nouns in TXC consist of a root and a null categorizer, and that the reduplication is a procedure where the null categorizer copies the phonological features of its single sister root. Based on this, a set of predictions are made concerning reduplication in TXC affixed nouns and attributive compound nouns, which will all be seen to be true.

3.1   Head Affixes

There are suffixes in TXC which are used to form nouns.

(10)   (a)   xiong   -bong
                countryside   -person
                'country bumpkin'

         (b)   rou   -dan
                meat   -person
                'blockhead'

The suffixes *-bong* and *-dan* have the meaning 'person who is associated with X', where X is the entity that is denoted by the item the suffix is merged with, similar to that of the English suffix *-er* in *teenager*, *foreigner* or *-y* in *fatty*. Both suffixes have pejorative connotation. (10a) denotes a kind of person, so categorial and semantic features of the suffix *-bong* project to the word *xiong-bong* which dominates the suffix *-bong*. Hence the suffix is

---

[4] There is some evidence that proper names in Chinese conform to the two-constituent condition. This is clearly not the case in all languages. We leave this issue for future research.

the head in (10a). Same in (10b), which denotes a kind of person, not a kind of meat, so the semantic and presumably the categorial features of the suffix *-dan* project to the resultant word *rou-dan*. So the suffix *-dan* is the head in (10b).

This means that the object-denoting items *xiong* 'countryside' and *rou* 'meat' that the suffixes in (10a) and (10b) are merged with are the non-head elements. Their status as non-heads is ensured if they are roots, not words, comparable to the roots in the Swedish derived nouns. As roots they have no categorial feature to project, and are hence by necessity non-heads. The structure of, for example (10b) would be (11).

(11)

```
          N
         / \
        R   n
        |   |
       rou -dan
```
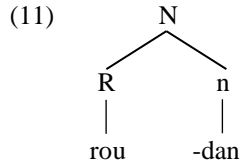
If *xiong* 'countryside' and *rou* 'meat' in (10a,b) were nouns, the head of the construct would not be determinable, and they would be predicted to be ill-formed.

In the last section we proposed that the reduplication in TXC nouns is a process where the null nominalizer copies the phonological matrix of its single sister root. In other words, the presence of the null nominalizer is crucial for the reduplication to take place in TXC nouns. Based on this and the classification of the object-denoting item *xiong* 'countryside' and *rou* 'meat', it is predicted that items such as *xiong* 'countryside' and *rou* 'meat' which are merged with the suffixes *-bong* and *–dan*, as in (12), cannot be reduplicated, in the absence a sister null nominalizer which would copy their phonological features. This prediction is true, as shown by the following examples:

(12) (a) *xiong   xiong   -bong
           countryside countryside -person

      (b) *rou   rou   -dan
           meat   meat   -person

As independent nouns, *xiong* 'countryside' and *rou* 'meat' can be, and have to be reduplicated. In this case they are merged with the null nominalizer, which will copy their phonological features in order to comply with the TXC version (9) of the two-constituent condition.

(13) (a) xiong   xiong
           countryside n
           'countryside'

      (b) rou   rou
           meat   n
           'meat'

The suffixes *–bong* and *–dan* themselves cannot be reduplicated, either.

(14) *xiong-bong-bong, *rou-dan-dan

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 features, by hypothesis just the semantic feature 'person' and a nominal feature.

This analysis of *-bong* and *-dan* is not obviously right. 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 (2010), Lowenstamm (2015), Creemers, Don, and Felser (2018). Lowenstamm (2015), for example, argues that the English derived noun *librarian* has basically the structure (15) using our notation.

(15)

```
            NP
           /  \
          /    n
         / \   |
        R   R  Ø
        |   |
     library -ian
```
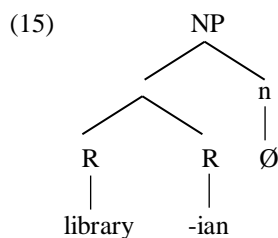
Under this view the interpretation of the word would not be compositionally derived – it could not be, as the [R,R] combination has no head – but would be acquired directly from the Encyclopedia. This would always be the case where two roots are merged to form a word (see also Zhang 2007, Bauke 2014: ch. 2, Hu and Perry 2017).

In section 5 we will argue that there are words in Mandarin and TXC that have this structure, including various kinds of non-compositional compounds. However, we do not adopt this analysis for words formed by the suffixes *–bong* and *–dan*. Classifying them as roots would require assuming that there is a special subcategory of roots which have a selection feature, selecting to merge with a root, and a linearization feature: they are always spelled out following their sister. We maintain that roots have no features other than semantic ones. But the derivational suffixes *-bong* and *–dan* have syntactic features: they are nominal and select a root. Being heads, they follow their sister: they are suffixes.

3.2   Non-Head Affixes

There are some suffixes in TXC which do not have effect on the category or the meaning of the resultant word:

(16)   (a)   mo   -e
              cat   -E
              'cat'

        (b)   za   -zi
              powder   -ZI
              'powder'

The suffix appears to have no effect on either the semantics or the category of the word: *mo-e* is a noun which means 'cat' and *za-zi* a noun which means 'powder'. Alternative forms are the reduplicated forms *mo mo* 'cat' and *za za* 'powder'. This suggests that the suffixes are devoid of any syntactic features, including categorial features; they would have a phonological matrix and nothing else. If so, the other constituent in (16a,b), *mo* and *za*, must be a noun, providing a head for the word. It cannot be a bare root, or the word would have no category. By hypothesis, this means that it is made up of a root and a null nominalizer. The structure of for example (16a) would be (17):

(17)
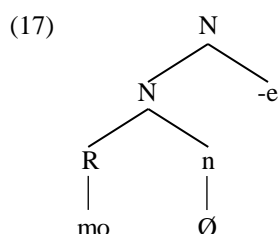
A prediction can be made based on this analysis of *mo* 'cat' and *za* 'powder' in (16a, b) and the procedure of reduplication in TXC nouns, which is that in the resultant affixed word, the item that the non-head suffix is merged with, can be reduplicated. This prediction is borne out:

(18)   (a)   mo   mo   -e
              cat   cat   -E
              'cat'

        (b)   za   za   -zi
              powder powder -ZI
              'powder'

Comparing (18) and (16), it can be seen that the reduplication is optional. This, we contend, is because condition (9) is satisfied already without reduplication, by the suffix. This means 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 condition (9) is not otherwise met, the reduplication must apply.

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

(19)   (a)   a-   yi
              A-   grandfather
              'grandfather'

(b)   ga-   chei
        GA-   bike
        'bike'

The prefix *a-* only has phonological features. With or without the prefix, (19a) denotes grandfather, which can be understood if the semantic and categorial features of the item *yi* project to the resultant word *a-yi*. Hence the item *yi* 'grandfather' is the head and the prefix *a-* is the non-head. Similarly in (19b), the prefix *ga-* is the non-head and *chei* 'bike' is the head, as features of *chei* 'bike' project to the resultant word *ga-chei* 'bike', while the prefix contributes nothing towards its interpretation. So *yi* 'grandfather' and *chei* 'bike' must be categories which are able to project. That is to say, they cannot be roots but must be nouns. That means they are made up of a root and a null nominalizer. This predicts that they can be reduplicated. This prediction turns out to be accurate, as (19a,b) can have the following reduplicated forms:

(20)   (a)   a-   yi   yi
                A-   grandfather   n
                'grandfather'

        (b)   ga-   chei   chei
                GA-   bike   n
                'bike'

The structure of, for example (19b) would be as follows:

(21)   N
        ┌───┴───┐
       ga-   N
            ┌───┴───┐
            R   n
            │   │
           chei   Ø
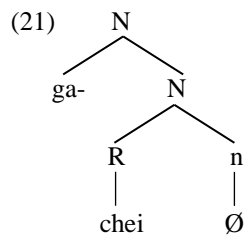
The null nominalizer may copy the phonological matrix of the root, optionally in this case, as condition (9) is satisfied anyway, by the prefix.

In addition to derivational affixes, there is also an inflectional affix in TXC nouns:

(22)   dueng   -men
        hole   -PLURAL
        'holes'

The pluralizing suffix *-men* is a syntactic category which will only merge with another syntactic category, that is with a noun (or possibly more correctly, NP); Li (1999), Ueda and Haraguchi (2008). A root cannot merge with an inflectional suffix. If the sister of the plural suffix is a noun, it will consist of a root and a null nominalizer, which predicts that it may undergo reduplication. This prediction is right as the following example shows:

(23)   dueng   dueng   -men
        hole   n   -PLURAL
        'holes'

The structure of (22) is (24):

(24)   N
            ┌───┴───┐
            N   -men
        ┌───┴───┐
        R   n
        │   │
      dueng   Ø
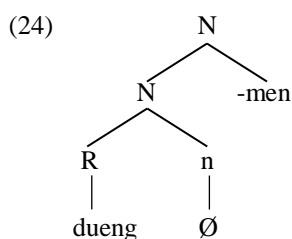
3.3 Attributive Compound Nouns in TXC

Above, based on how reduplication operates in TXC and the analysis of components of affixed nouns, predictions which are made are confirmed. Furthermore, based on the same principles, a prediction can be made concerning attributive compounds, which is also accurate:

(25) (a)   cei   mo
                   vegetable   steamed bun
                   'vegetable   steamed bun'

         (b)   mei   hu
                   ink   box
                   'ink box'

Based on the interpretation, *mo* 'steamed bun' in (25a) and *hu* 'box' (25b) is the head and *cei* 'vegetable' in (25a) and *mei* 'ink' in (25b) is the modifier. Thus *mo* 'steamed bun' in (25a) and *hu* 'box' in (25b) will be nouns, not bare roots, and as such they are made up of a root and a null nominalizer. Hence they are predicted to allow reduplication. This prediction is shown to bear out in the following reduplicated forms:

(26) (a)   cei   mo   mo
                   vegetable   steamed bun   n
                   'vegetable steamed bun'

         (b)   mei   hu   hu
                   ink   box   n
                   'ink box'

On the other hand, *cei* 'vegetable' in (25a) and *mei* 'ink' in (25b), which are the non-heads, cannot be nouns that project, made up of a root and a null nominalizer. Instead, following the theory in section 2, they are bare roots. Since reduplication in TXC nouns requires a null nominalizer, it is predicted, for the non-head in the attributive compound in TXC, that it cannot be reduplicated. This prediction is confirmed by the following ungrammatical examples:

(27) (a)   *cei   cei   mo
                   vegetable   n   steamed bun

         (b)   *mei   mei   hu
                   ink   n   box

As (25a,b) are both grammatical, reduplication of the head in attributive compounds is apparently optional. Again, this is predicted if the reduplication is dependent on condition (9): In an attributive compound with two pronounced components this condition is satisfied already without reduplication, hence reduplication is allowed but not required.

      We may conclude that the distribution of reduplication in TXC nouns is predicted on the basis of the morphosyntactic properties of the components of the nouns, namely, whether the component is a head or not. Whether a word component is a head or not is not phonologically marked, in TXC. This implies that a phonological approach could not explain the distribution of reduplication in TXC nouns discussed in this section. We thus have another reason to reject the alternative idea that the reduplication of nouns in TXC is phonologically motivated.

4. Bound words

4.1 Bound words in Mandarin

Chinese has a class of content words distinguished by the property that they have to be morphologically bound. The following are some examples of such lexical items from Mandarin, exemplifying nouns, adjectives, and verbs.

(28)   *nao* 'brain'   *dian* 'dictionary'   *ying* 'film'   *gui* 'rule'   *yi* 'chair'   *shi* 'stone'
       piao 'pretty'   *cai* 'colourful'   *mian* 'shy'   *xing* 'walk' *yun* 'carry' *hui* 'return'

None of these lexical items can stand alone as a free word in a phrase or a sentence. For reasons of space, we exemplify this with just one word, the noun *yi* 'chair'.

(29)   *yi   ba   yi   (Mandarin)
        a   CL   chair
        Intended reading 'a chair'

(30) shows that *yi* can occur as part of a compound, while (31) shows that it can occur as a free word if it is merged with an affix. This is true of all of the items in (28).

(30) yi   ba   chang   yi
      a   CL   long   chair
      'a long chair'

(31) yi   ba   yi   -zi
      a   CL   chair   -ZI
      'a chair'

In the literature these items are called bound stems (Dai 1992:40,75-76)or bound roots (Sproat and Shih 1997, 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, hence there are no bound stems, and it does not make sense to classify them as 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. They are content words, but unlike free content words, they are words with inherent word category. In other words, they are not a combination of a root and a categorizer but lack internal structure. Bound words are like functional heads in this regard.

There are similarities between the bound word and the root, in Mandarin and TXC. First of all, the bound word and the root both have lexical content. Further, neither of them can stand alone as a free content word in a phrase; both need to merge with another item to form a free content word. However, the bound word is crucially different from a root, in that a root can merge with an item which is not pronounced, i.e. a null categorizer, and together they can form a free content word. A bound word cannot do this. Following Wang (2018) we claim that this is because it has inherent category, it is a single morphological item with semantic features and a syntactic categorial feature, and thus merging with a null categorizer is excluded on the grounds of economy. However, following the proposed two-constituent condition (8) in section 2.2, which says that a free content word minimally contains two constituents, a bound word has to merge with another constituent, to form a free word.

So bound words, unlike free content words, are not made up of a root merged with a categorizer. This makes a prediction: Since reduplication in TXC nouns is derived by copying the phonological matrix of a root onto a sister null nominalizer, bound words in TXC cannot be reduplicated. The prediction is true, as we will now show.

The following is a list of bound words in TXC:

(32) *yi* 'clothing'   *ta* 'inner shirt'   *ti* 'drawer'   *can* 'shovel'   *gei* 'armpit'
      jieng 'towel'   *nong* 'dirty'   *zuen* 'beautiful'   *ji* 'solid'

(33a) illustrates the fact that *can* 'shovel' cannot be used alone as head of a phrase in TXC, but can be, if it is merged with another word or root in a compound, as in (33b), or if it is merged with an affix, as in (33c).

(33) (a) *qieng   zi   can
           light   DE   shovel
           Intended reading: light shovel

      (b) qieng   zi   mu   can
           light   DE   wood   shovel

      (c) qieng   zi   can   -zi
           light   DE   shovel -ZI

Like its counterpart in Mandarin, the bound word in TXC is, by hypothesis, a single item with a categorial feature. In other words, the bound word is not composed of a root and a null categorizer, and does not have internal structure. (34) demonstrates that the reduplicated form of *can* 'shovel' is ungrammatical as head of an NP.

(34) *qieng   zi   can   can
       light   DE   shovel shovel
       Intended reading: 'light shovel'

In fact, reduplication of the bound word is ungrammatical in any context. (35) shows that it is ungrammatical when the bound word is merged with a root, and thus functions as head of a word; recall from section 3 that reduplication of the head of a compound noun is otherwise optional.

(35) (a) *shou   jieng   jieng
          hand   towel   n

(36) shows that reduplication is also ungrammatical when the bound word is merged with an acategorial affix; recall that this is a context where reduplication is otherwise optional in TXC.[5]

(36) (b) *jieng   jieng   -e
          towel   towel   -E

5.   Non-compositional compounds

The compounds discussed so far have had compositional structure and semantics: The meaning of the compound has been predictable from the meaning of the constituents and how they are merged. Not all compounds are like this. There are various kinds of non-compositional compounds. Space does not allow a detailed exposition here, but we will demonstrate that our analysis of reduplication in TXC makes the right prediction for such compounds in TXC.

Various types of non-compositional compounds in Chinese have been discussed by Zhang (2007) (for Mandarin), Hu and Perry (2017) (for Yixing Chinese), and Wang (2018) (for Mandarin and TXC). They all present arguments that the non-compositional compounds are derived by 'root merger', that is by merging two categoryless roots, followed by merge of a null categorizer (see also Bauke 2014: ch.2). The structure of a non-compositional compound noun would thus be (37):

(37)

```
              N
            /   \
          /       \
        /           n
      /   \
    R       R
```
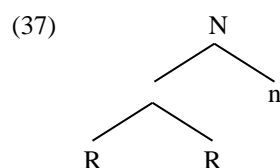
Such compounds would not have a compositionally derived interpretation as the unit of two merged roots has no head and no label (in the absence of any syntactic features). For the same reason they cannot have a linear order derived by the regular rule(s). Instead, their interpretation and linear order comes directly from the Lexicon (or Encyclopedia), where the compounds are listed.

One example of a non-compositional compound is the type called parallel compounds in some of the literature, including Chao (1968) and Wang (2018). They consist of two items which are associated semantically, but the association can be quite loose and is sometimes quite opaque; they are 'parallel':

(38) (a) nian   wu
          eye   pit
          'eye'

      (b) pi   mo
          leather   fur
          'fur'

Even though the meaning of the compound is (roughly) equivalent to the meaning of one of the constituents, we do not want to say that that constituent is the head and the other constituent is a modifier. There is no modification relation, or any other well-defined relation between the constituents. Note also that they do not follow the general rule for word-internal order (the 'Right-hand Head Rule'), or indeed any other rule. In (38a) the initial constituent has a meaning equivalent to the meaning of the compound, in (38b) it is the final constituent. Instead, they are best regarded as a subtype of coordinative compounds. Following Wang (2018) we propose that they are derived by root merger. The structure of, for example, (38a) is (39):

---

[5] Please see section 3.2 for the optional reduplication in affixed words in TXC.

The fact that two bound words can merge and form a compound with compositional semantics has interesting theoretical consequences, not discussed here for reasons of space; see Wang (2018).

(39)

```
                N
              /   \
            /       n
          /   \
         R     R
         |     |
        nian   wu
        eye    pit
```
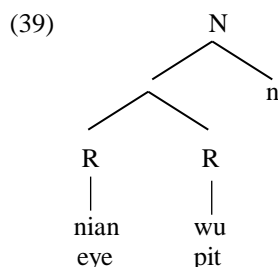
Two roots are merged, forming an unlabelled, categoryless unit, which is merged with a null nominalizer. The meaning is not derived compositionally, and the linear order is not derived by rule, but instead meaning and spelled-out form are both drawn directly from the Encyclopedia.

The prediction now is that the constituents in this type of compound cannot be reduplicated. The compound has a null nominalizer, but as it is not the sister of either of the roots, reduplication will not apply. The prediction is right.

(40) (a)   *nian nian wu
      (b)   *nian wu wu

As free nouns, *nian* and *wu* have to be reduplicated, and when occurring as heads of compositional compounds they can be reduplicated, but as constituents in a non-compositional compound they cannot, as our theory would predict.[6] The generalization holds true of non-compositional compounds in general in TXC: The constituents cannot be reduplicated, as predicted if they are derived by root merger.

6. Conclusions

In Traditional Xining Chinese (TXC) free nouns are always reduplicated. We claim that this is because (a) free content words consist of a root and a categorizer, (b) there is a condition on content words that they must contain at least two constituents, which may be a root and a categorizer, (c) TXC has a version of this condition which says that the two constituents, in the case of nouns, must be pronounced. If this condition is not already satisfied, as in the case of a compound or a root merged with an overt affix, the condition is satisfied by reduplication: the nominalizer copies the phonological matrix of the sister root. We have shown that this hypothesis makes the right predictions for all kinds of complex words in the language.

References

Bauke, Leah.S. 2014. *Symmetry breaking in syntax and the lexicon*. Amsterdam Philadelphia: John Benjamins Publishing Company.

Bell, Daniel. 2016. 'Syntactic change in Xining Mandarin'. PhD dissertation, Newcastle University.

Cardinaletti, Anna, and Michael Starke 2000. *"Overview: the grammar (and acquisition) of clitics."* In *The acquisition of scrambling and cliticization,* edited by Powers, Susan and Cornelia Hamann, 165–186. Springer: Dordrecht.

Chao, Yuan. Ren. 1968. *The grammar of spoken Chinese*. Berkeley: University of California Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Creemers, Ava., Jan, Don and Fenger Paula. 2018. "Some affixes are roots, others are heads." *Natural Language and Linguistic Theory* 36: 45-84.

Dai, Xiang-Ling. 1992. "Chinese morphology and its interface with the syntax." PhD dissertation, Ohio State University.

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 dissertation, Utrecht University.

Déchaine, Rose-Marie and Martina Wiltschko.2002. "Decomposing pronouns." *Linguistic Inquiry* 3:409-422.

Embick, David and Rolf Noyer. 2007. "Distributed morphology and the syntax/morphology interface." In *The oxford handbook of linguistic interfaces*, edited by 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.

Hall, Alan. 1999. "The phonological word: a review." In *Studies on the phonological word*, edited by Tacy Alan Hall and Ursula Kleinhenz, 1-22. Amsterdam: John Benjamins Publishing Company.

---

[6] The constituents of non-compositional compounds can be bound words. As a bound word can label a phrase (in some contexts), this could be seen as possibly affecting the derivation. We leave this complication for future work. Reduplication is ruled out in any case, as bound words are never reduplicated.

Harley, Heidi and Noyer, R. 1999. "Distributed morphology." *Glot International* 4:3-9.

Harley, Heidi. 2011. 'Compounding in distributed morphology'. In *The oxford handbook of compounding*, edited by Rochelle Lieber and Pavol Štekauer, 129-144. Oxford: Oxford University Press.

Harris, James. 1991. "The exponence of gender in Spanish." *Linguistic Inquiry* 22: 27-62.

Holmberg, Anders. 1992. "Properties of non-heads in compounds: a case study." *Working Paper in Scandinavian Syntax* 49:27-58.

Hu, Xu and J.Joseph, Perry.2017."The syntax and phonology of non-compositional compounds in Yixing Chinese." *Natural Language and Linguistic Theory* 36:701. Accessed August 14, 2018. DOI: http://dx.doi.org/10.1007/s11049-017-9386-8

Josefsson, Gunlög. 1997. *On the principles of word formation in Swedish*. Kent: Chartwel-Bratt Ltd.

Josefsson, Gunlög. 1998. *Minimal words in a minimal syntax*. Amsterdam: John Benjamins

Kiefer Ference. 1970. *Swedish morphology*. Stockholm: Skriptor.

Li, Yen-fei, Audry. 1999. "Plurality in a classifier language." *Journal of East Asian Linguistics*. 8:75-99

Lowenstamm, Jean. 2015. "Derivational affixes as roots: phrasal spell-out meets English stress-shift." In *The syntax of roots and the roots of syntax,* edited by Artemis Alexiadou, Hagit Borer and Florian Schäfer, 230-259. Oxford University Press.

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.

McCarthy, John, and Alan Prince. 1990. "Foot and word in prosodic morphology: the Arabic broken plurals." *Natural Language and Linguistic Theory* 8:209-282.

Mukai, Makiko. 2008. "Recursive compounds." *Word Structure* 1: 178-198.

Mukai, Makiko. 2018. *A comparative study of compound words*. Tokyo: Hituzi Syobo Publishing.

Packard, Jerome. 2000. *The morphology of Chinese: a linguistic and cognitive approach.* Cambridge: Cambridge University Press.

Pirani, Laura. 2008. "Bound roots in Mandarin Chinese and comparison with European semi-words." *Proceedings of the 20th North American Conference on Chinese Linguistics*. 1:261-277.

Ren, Bi.Sheng. 2006. *Qinghai fangyan yufa zhuanti yanjiu [The grammar of the Qinghai dialect].* Xining: Qinghai renmin chubanshe *[Qinghai People Press]*

Sproat, Richard. and Shih, Chilin 1996. "A corpus-based analysis of Mandarin nominal root compound." *Journal of East Asian Linguistics* 5: 49-71.

Travis, Lisa. 2001. "The syntax of reduplication." *Proceedings of Northeast Linguistics Society* (NELS) 31, GLSA, 454-469.

Ueda, Yasuki and Tomoko Hamaguchi. 2008. "Plurality in Japanese and Chinese." *Nanzan Linguistics Special Issue* 229-242.

Wang, Qi. 2018. The structure of nouns in Old Xining and Modern Standard Chinese. PhD dissertation, Newcastle University.

Wang, Shuang.Cheng. 2009. "Xining fangyan chongdieshi." [Reduplication in Xining Chinese]. *Qinghai shifan daxue minzu shifan xueyuan xuebao* 5: 1-4. [Qinghai Normal University Minzu Department Academic Newspaper**]**

Williams, Edwin, 1981. "On the notions lexically related and head of a word." *Linguistic Inquiry* 12:245-274.

Zhang, Niina.Ning. 2007. "Root merger in Chinese compounds." *Studia Linguistica* 2: 170-184.