The structural sources for syllabic restrictions in Mandarin*

Lisa Lai-Shen Cheng\textsuperscript{a} & Hang Cheng\textsuperscript{b}

\textsuperscript{a}Leiden University Centre for Linguistics
\textsuperscript{b}Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies

1. Introduction

\textsuperscript{a}Lu (1981), \textsuperscript{a}Fu (1994), and \textsuperscript{b}Deng (2021) among many others note that there is a syllabic restriction on nominalization in Mandarin Chinese, namely that only disyllabic verbs can be nominalized. Consider the contrast indicated in (1). All the verbs in bold can be independently used as a verb. However, only the disyllabic verbs can occur in the post-	extit{de} position.

\begin{enumerate}
\item \textit{tāmen de bào/gào\textsuperscript{4}}
\begin{enumerate}
\item 3pl \textit{de} \textit{report/report} ‘their report’
\end{enumerate}
\item \textit{tā duì biérén de gōngjī/jī}
\begin{enumerate}
\item 3sg \textit{prep} other.people \textit{de} \textit{attack/attack} ‘the attack on other people’
\end{enumerate}
\end{enumerate}

On the other hand, there appears to be a “reversed” syllabic restriction. Even though Chinese languages have very few denominal verbs (see \textsuperscript{c}Chan and Tai (1995) and \textsuperscript{c}Yuan and Jiang (2018)) in comparison with a language like English, only monosyllabic nouns can be denominals. \textsuperscript{c}Tsai (2016) cites a couple of examples in modern colloquial Mandarin, as in (2) (from Tsai 2016, ex. 4a,b). Importantly, \textit{diàn} is the first syllable of \textit{diànhuà} ‘telephone’ while \textit{duǎn} is the first syllable of \textit{duǎnxìn} ‘text’. The full disyllabic verbs cannot replace the monosyllabic ones.

\begin{enumerate}
\item \textit{Yǒu shénme wèntí, qǐng diàn/diànhuà fúwùchù.}
\begin{enumerate}
\item have what \textit{question} please \textit{phone} reception ‘If you have any question, please call the reception.’
\end{enumerate}
\end{enumerate}

\textsuperscript{a}We are happy to dedicate this paper to Hagit. She has been a great inspiration to us. Lisa would also like to thank Hagit for all the wonderful discussions including topics outside of linguistics.

\textsuperscript{b}The current paper confines our discussion about “nominalization” to the X-	extit{de}-Y construction and a couple of other nominal constructions that observe the same syllabic restrictions. Other cases where a common verbal constituent appearing in positions that are typically reserved for nominal constituents, such as the subject position (see (i)), will not be discussed as a similar syllabic restriction is not observed.

\textsuperscript{c}The following abbreviations are used in the glosses: 1 = first person, 2 = second person, 3 = third person, \textit{ba} = \textit{bǎ}, \textit{clf} = classifier, \textit{cop} = copula, \textit{de} = \textit{de}, \textit{dem} = demonstrative, \textit{exp} = experiential, \textit{imp} = imperative, \textit{neg} = negative, \textit{pl} = plural, \textit{poss} = possessive, \textit{prep} = preposition, \textit{prf} = perfect, \textit{sg} = singular.
The contrasting syllabic restrictions are seemingly compatible with Duanmu’s (2007:183) observations of the relation between word length, word categories and syntactic function. He notes that in Mandarin syntactic heads have a high percentage of monosyllables while nonheads tend to have a low percentage of monosyllables. Following this, it is not entirely surprising that nominalized constituents in non-head positions are disyllabic while denominal verbs (i.e., heads) are monosyllabic. However, Duanmu’s generalization is not sufficient to account for all the facts regarding the syllabic restrictions in question. For one thing, as Duanmu notes, though being heads, verbs generally can be disyllabic in Mandarin. Similarly, as we will show later, the post-*de* position, being non-heads, does not always require disyllabic constituents. Hence, it is unclear what rules out disyllabic denominal verbs or monosyllabic nominalization. For the other, as will be discussed below, both nominalization and denominal verbs are subject to other restrictions as well.

To understand the syllabic restriction on nominalization, Fu (1994) suggests two possible lines of inquiry, but does not commit to one of them. The first one is a phonological restriction similar to English comparative *-er*. The second line of inquiry concerns derived structures: disyllabic items are necessarily derived since each Mandarin syllable corresponds to a morpheme. As far as we know, there is no discussion in the literature concerning the restrictions on denominal verbs.

In this paper, we argue against both a purely phonological account and a prosodic account for both syllabic restrictions. Instead, we argue that the sources of the syllabic restrictions in both nominalization and denominal verbs are structural. The next section provides a more elaborate description of syllabic restrictions on Mandarin nominalization and denominal verbs. In section 3, we put forth our account of the syllabic restrictions on nominalization. Section 4 presents an analysis of the syllabic restriction in denominals by considering the possible type(s) of denominals. In section 5, we conclude the paper by making a comparison between the Exoskeletal framework and the Distributive Morphology framework in nominalization.

2. The phenomena

2.1. Background: The long-short word pairs in Mandarin

It has been observed that, in Mandarin, a large amount of words (or morphemes) have elastic length (Li 1998, Duanmu 2012, 2016). That is, such words (or morphemes) can “be long (disyllabic) or short (monosyllabic), with more or less similar meanings” (Duanmu 2012, 89) and “where the former contains the latter” (Duanmu 2016, 3). For instance, as shown in (3), the disyllabic words in the first column have a similar meaning as the monosyllabic ones in the second column. Moreover, the disyllabic words consist of the monosyllabic counterpart (or morpheme) and an extra element.

<table>
<thead>
<tr>
<th>Disyllabic</th>
<th>Monosyllabic</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. huǐmiè</td>
<td>huǐ</td>
<td>‘to destroy’</td>
</tr>
<tr>
<td>b. diàochá</td>
<td>chá</td>
<td>‘to investigate’</td>
</tr>
<tr>
<td>c. méitàn</td>
<td>méi</td>
<td>‘coal’</td>
</tr>
<tr>
<td>d. túsúshū</td>
<td>shū</td>
<td>‘book’</td>
</tr>
</tbody>
</table>

Note that, the long-short word pairs (or “dual vocabulary”, another term that Duanmu uses)

It is well-known that there are a number of monomorphemic disyllabic words in Mandarin, such as *pútao* ‘grape’ and *bōli* ‘glass’. See for instance Fu (1994: 54) among many others.
can be found in both verbs (e.g., (3a) and (3b)) and nouns (e.g., (3c) and (3d)). Also, the monosyllabic words can correspond to either the first (e.g., (3a) and (3c)) or the second (e.g., (3b) and (3d)) element of the disyllabic words. Nonetheless, in many cases other than those that are listed in (3), the long and short correspondences do have interpretational distinctions, and especially on collocation. For instance, you can either chá or diàochá a case, both of which mean ‘to investigate a case’. You can also chá a dictionary, which means ‘to look up a dictionary’; however, you cannot diàochá a dictionary (unless it concerns physical evidence of a case involving a dictionary). Apart from the interpretation and collocation distinctions, word length choices are not always free. The current paper will not investigate all the restrictions on word length choices but concentrate on the above-mentioned two contexts – (a) nominalization which only allows for disyllabic verbs, and (b) denominal verbs which can only be monosyllabic.

2.2. Restrictions on Mandarin de-Nominalization

Fu (1994) notes that nominalization in Mandarin is subject to a syllabic restriction and some semantic restrictions. Consider first a couple of examples of nominalization (from Fu 1994: 53, with adaptations for the b-examples).

(4) a. tāmen de tánhuà/*tán
   3PL POSS talk
   ‘their conversation’

b. Tāmen tánhuà/tán le.
   3PL talk PRF
   ‘They talked.’

(5) a. Zhāngsān dui zhèi-jìàn shì de diàochá/*chá
   Zhangsan PREP DEM-CLF matter DE investigate
   ‘Zhangsan’s investigation of this matter’

b. Zhāngsān shàng-ge yuè diàochá/chá le zhè-jiàn shì.
   Zhangsan last-CLF month investigate PRF DEM-CLF matter
   ‘Zhangsan investigated this matter last month.’

As shown in (4) and (5), it is possible to nominalize tánhuà ‘to talk’ and diàochá ‘to investigate’ but not tán and chá, which also mean ‘to talk’ and ‘to investigate’, respectively. Clearly, monosyllabic verbs are excluded from this type of nominalization. Note that, nominalization is not restricted to disyllabic verbs. Verbs that have more than two syllables can also be nominalized, as shown in (6).

(6) a. shíwù de zàishēngchǎn
   food DE reproduce
   ‘reproduction of food’

b. liǎng-ge xiǎoshí de yù-jiāgōng
   two-CLF hour DE pre-process
   ‘two hours’ pre-processing’

However, not all the di- and multi-syllabic verbs can be nominalized. First, with few exceptions, nominalization of disyllabic monomorphemic words are rare. Admittedly, examples like (7a) are attested in corpora, though the amount of such cases is quite small in comparison to their verbal counterparts like (7b).

For instance, in BCC corpus (http://bcc.blcu.edu.cn/), a corpus for contemporary and ancient Chinese which contains more than 15 billion words, there are around a hundred entries of nominal uses of zhànzhuǎn, in comparison to more than six thousand entries of verbal uses of it.
Moreover, as noted in Fu (1994), resultatives such as dā-shāng ‘hit-wound’ cannot be nominalized, even though they satisfy the disyllabic restriction. Deng (2021) claims that all verbs that can be nominalized by de are formed by coordinating two morphemes that have similar meanings (if not strictly identical). Take diàochá ‘to investigate’ as an example. Both morphemes, diào and chá, roughly mean ‘to investigate’, similar to their combined meaning.

Although Deng (2021) treats all that he labels as “event nominals” as conjunction of two morphemes, we suggest that at least some of these verbs involve a modification relation between the two composing morphemes. Below are two lists of verbs that can be nominalized; and they also qualify as “event nominals”. All the verbs in the first group contain biān ‘create, edit’ and those in the second group contain lùn ‘discuss’. Although both biān and lùn are closely related to the other element in each verb, they are by no means synonymous with the other morphemes.

For (8a), biān modifies different types of activities. Similarly, for (8b), various types of lùn are depicted.

### Examples

#### a. biān + X
- biān-dǎo ‘to write and direct (a play)’
- biān-jí ‘to edit’
- biān-shěn ‘to review’
- biān-yì ‘to translate’

#### b. X + lùn
- tǎo-lùn ‘to discuss’
- biàn-lùn ‘to debate’
- zhēng-lùn ‘to argue’
- píng-lùn ‘to comment’
- tán-lùn ‘to talk about’

There are two more properties of de-nominalization that we would like to highlight here before returning to Mandarin denominal verbs. The first concerns the basics of the marker de, which is a crucial component in nominalization. It has been called the nominalization marker by Paris (1979) among many others. As is well-known, de is used in nominalization, relativization and basically in all nominal modification, as we can see in (9b) and (10).

#### a. Tāmen tāolùn-guo nà-ge wèntí.
- 3PL discuss-exp dem-clf question

‘They have discussed that question.’

#### b. tāmen tāolùn de nà-ge wèntí
- 3PL discuss-de dem-clf question

‘the question that they discuss(ed)’

(5) In the traditional approach of Chinese grammar analysis, conjunction at the word-level commonly concerns two morphemes that can be taken as synonyms or antonyms. An example for the former is tāo-lùn ‘to discuss’. These two morphemes, tāo and lùn, are synonyms, roughly with the meaning ‘to discuss’. An example for the latter is zǎo-wǎn ‘sooner or later’. zǎo means ‘early, morning’ and wǎn means ‘late, evening’, and they are antonyms. Crucially, words like diàn-huà [electric-speech] ‘telephone’, in which the first morpheme supposedly modifies the second one, are analyzed as compounding in a different manner – piānzhéng [side-center], namely, modification.
Fu (1994) argues that *de* is phonologically attached to the phrase/clause preceding it (i.e., as a (phonological) suffix), by showing that *de* can be stranded. One may thus hypothesize that the syllabic restriction we see in nominalization has to do with a restriction concerning the elements following *de*, i.e., only disyllabic elements can follow *de*. However, as we see in (10b-d), the constituent following *de* can be monosyllabic. This means that the syllabic restriction under discussion is not a pure phonological restriction; and (10d) confirms that this is also not the case when we are dealing with a (longer) clausal element (i.e., a relative clause) before *de*.

The second property to highlight concerns basic noun phrase structure in Chinese languages. As is well-known, Chinese languages have classifiers. One property that is less discussed is the word order of the dem(onstrative)-cl(assifier) sequence and the noun. Fu (1994) and most of the *de*-nominalization literature use bare nouns (i.e., without Dem-Cl sequence). As we can see in (9b), following *de*, we can have Dem-Cl-N. However, it is also possible to have the order illustrated in (11), where the *de*-phrase intervenes between the Dem-Cl sequence and the noun.

(11) nà-ge [tāmén tǎolùn de] wèntí
      DEM-CLF 3PL discuss DE question
      ‘the question that they discuss(ed)’

Fu (1994) argues that classifiers can display the interpretation of the nominalized elements, as illustrated by the contrast between the examples in (12) (from Fu 1994, p.66, ex. (33)).

---

(i) a. Zhāngsān de lǎn hùi hài le tā.
       Zhangsan de lazy will harm 3SG
       ‘Zhangsan’s laziness will ruin him.’

b. Lǐsì de qínfèn yǒumùgòngdǔ.
   Lisi de diligent obvious
   ‘Lisi’s diligence is widely recognized.’
(12) a. Tā duì fāngyán de nèi-çi/*piān diàochá jìnxíng le sān-ge yuè
   3SG PREP dialect DE that-CLF/clf investigation proceed PRF three-CLF month
   ‘That investigation of the dialect by him went on for three months.’

b. Tā de nèi-çi/piān diàochá fābǐāo le.
   3SG DE that-CLF/clf investigation publish PRF
   ‘His investigation got published.’

Ci indicates frequency or occurrence, while the classifier piān is used for articles and reports. In (12a), only ci can be used because of the matrix verb jìnxíng ‘proceed’ requires a nominal which can be interpreted as a process nominal (see also Deng 2017 for a light verb analysis of jìnxíng). In contrast, in (12b), the matrix verb fābǐāo ‘publish’ is not compatible with a process nominal. Fu further suggests that when we have an argument in the nominalization (i.e., fāngyán ‘dialect’ in (12a)), we are dealing with a process nominal, and only the classifier ci can be used. More on the use of classifiers will be addressed in section 3.3.

2.3. Restrictions on Mandarin denominal verbs

There are two remarkable differences concerning denominal verbs between English and Chinese languages. First, in Chinese languages, there is no denominal verbs of the kind such as sweat, dance, sneeze, as discussed in Harley (2005), rather, all these counterparts are V+N combinations as in (13) (see also Cheng and Sybesma 1998).

(13) a. liú-hàn
   flow-sweat
   ‘to sweat’

b. tiào-wǔ
   jump-dance
   ‘to dance’

c. dǎ-pēntì
   hit-sneeze
   ‘to sneeze’

Bisang (2008) discusses the “underspecification” of nouns and verbs in Late Archaic Chinese, which represents the late stage of Old Chinese (11th – 3rd centuries BC). He argues that about 50% of the morphemes are not specified for a particular word class. However, Yuan and Jiang (2018) show that though “the phenomenon of nouns surfacing as verbs is highly prevalent in Old Chinese” (p.129), the number of denominal verbs remain quite small: they found 132 denominal verbs, which correspond to 5.4% of the words that were frequently used in pre-Qin texts. Furthermore, they indicate that the number of denominal verbs decreases throughout the history of Chinese. Chan and Tai (1995) show that only 110 denominal verbs can be identified in Mandarin (a similar size for Cantonese and Taiwanese). In comparison, Clark and Clark (1979) identify more than 1300 denominal verbs in English.

In addition, Clark and Clark (1979) sort the denominal verbs they identified into nine categories, namely, locatum verbs, location verbs, duration verbs, agent verbs, experiencer verbs, goal verbs, source verbs, instrumental verbs, and a miscellaneous class. Chan and Tai (1995) show that denominal verbs in the Mandarin corpus fall under four out of nine categories – instrument, goal, location, and locatum. Among the four classes of Mandarin denominal verbs, instrument verbs are the most common, while locatum and location verbs are rare. Yuan and Jiang (2018) confirm this picture, comparing denominal verbs in Old Chinese, Modern Chinese, and English. (14) exemplifies denominal verbs of each class in Mandarin.
Four classes of denominals in Mandarin

- **chuí** ‘hammer’ Instrument
- **duī** ‘heap’ Goal
- **dǐng** ‘top’ Location
- **qī** ‘paint’ Locatum

It should be noted that the syllabic restriction of Mandarin denominal verbs has not received much attention. Most of the denominal verbs the previous studies have identified are monosyllabic. More crucially, as we mentioned in the beginning of this paper, only the monosyllabic words of the long-short word pairs can be denominalized, if other requirements are satisfied. The most representative example is repeated below as (15).

(15) **Yǒu** shénme wèntí, qǐng diàn/**diànhuà** fúwūchù.
    have what question please phone reception
    ‘If you have any question, please call the reception.’

Note that a more common way to express ‘make a phone call’ in Mandarin (instead of using a denominal verb) is to insert an extra verbal element **dǎ** ‘to hit’.

(16) a. **Yǒu** shénme wèntí, qǐng gěi fúwūchù **dǎ** diànhuà.
    have what question please give reception hit phone
    ‘If you have any question, please call the reception.’

b. **Wǒ** gěi Lǐsì **dǎ** hit le sān-ge diànhuà.
    1sg give Lisi hit prf three-CLF phone
    ‘I called Lisi three times.’

However, unlike the denominal verb **diàn**, **dǎ**-**diànhuà** cannot directly take the goal as its complement, indicated by the ungrammaticality of the sentences in (17) (cf. [15] and (16a)). We argue in Cheng and Cheng (2021) that **dǎ** ‘hit’ in Mandarin is a verbalizer, among several other elements used to be identified as light verbs. This suggests that the structure permitting **diàn** (e.g., [15]) differs from the verbalized constructions such as **dǎ**-**diànhuà**.

(17) a. *Yǒu** shénme wèntí, qǐng **dǎ** diànhuà fúwūchù.
    have what question please hit phone reception
b. *Wǒ **dǎ** hit le sān-ge diànhuà Lǐsì.
    1sg hit prf three-CLF phone Lisi

2.4. Interim summary

The above sections delineate the restrictions on nominalization (or “derived nouns”, see Fu 1994 and Fu et al. 2001) and denominal verbs in Mandarin Chinese. Both processes are subject to syllabic restrictions and semantic restrictions. With respect to the syllabic restrictions, de-nominalization excludes monosyllabic verbs while denominal verbs can only permit monosyllables. Nominalization and denominal verbs are subject to certain semantic restrictions, too. For instance, resultatives cannot be nominalized. On the other hand, the majority of denominal verbs in Mandarin denote instruments. The differences between Mandarin nominalization and denominal verbs can be summarized in Table 1.

---

See Fu (1994) a possible explanation of why resultatives cannot be nominalized.
Table 1: Restrictions on Mandarin nominalization and Denominal verbs

<table>
<thead>
<tr>
<th></th>
<th>Nominalization (derived nouns)</th>
<th>Denominal verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syllabic restrictions</td>
<td>disyllabic</td>
<td>monosyllabic</td>
</tr>
<tr>
<td>Semantic constraints</td>
<td>no resultatives</td>
<td>mainly instrumental</td>
</tr>
</tbody>
</table>

In the next sections, we argue that although a phonological approach (such as the Non-head stress rules in Duanmu (2007)) can partially account for the syllabic restrictions in question, the restrictions are essentially syntactic. Nominalization requires a derived and categorized structure, whereas denominal verbs are closely related to a restriction which excludes roots other than instrumental roots.

3. The syllabic restriction on nominalization

We have argued briefly in section 2.2 that syllabic restriction on nominalization cannot be a pure phonological restriction simply because the post-de position allows for monosyllabic nouns. In this section, we first provide a more elaborate discussion from the metrical perspective and show that it is also insufficient as an explanation of the disyllabic restriction. We will provide more data to argue that an account based on derived structures can provide us with a more satisfactory account for a disyllabic requirement.

3.1. A prosodic account is insufficient

In addition to Fu’s (1994) phonological proposal that the syllabic restriction on Mandarin nominalization may be similar to English comparative -er, Duanmu (2007) also advances a prosodic account for such a restriction. He considers the disyllabic restriction in nominalization to fall under the flexible word-length problem. Specifically, Duanmu argues that variation in word length is governed by stress. In the [X de Y] nominalization structure, both X and Y are non-heads, and following the Nonhead stress rule (18), they must therefore receive phrasal stress.

(18) Nonhead stress
In the syntactic structure [X XP] (or [XP X]), where X is the syntactic head and XP the syntactic nonhead, XP should be stressed.

Furthermore, given the interaction between (19) and (20), Duanmu (2007) argues that if a disyllabic form is available, a word will choose the disyllabic form in the position of Y in [X de Y] nominalization structures.

(19) Stress length
Phrasal stress should be carried by a syllabic foot.
This can be satisfied by two syllables (SS) or by a syllable followed by an empty beat (S∅).\[^{10}\]

(20) Anti-allomorphy
A stressed word should keep the same phonological shape. (If a word has a disyllabic shape, it should be used when the word has phrasal stress.)

Consider diàochá and chá in (5) as an example: both have the meaning ‘to investigate’. Under Duanmu’s analysis, if chá is put in the Y position in [X de Y] nominalization, the Nonhead stress rule in (18) would require that chá carries phrasal stress. Given Anti-allomorphy in (20), since Duanmu (2007) notes that the empty beat can only occur if it has a phonetic correlate, such as a pause or the lengthening of the preceding syllable.

\[^{10}\]Duanmu (2007) notes that the empty beat can only occur if it has a phonetic correlate, such as a pause or the lengthening of the preceding syllable.
chá and diàochá are synonymous, diàochá would be used in the nominalization. According to Duanmu, this is why we have disyllabic words in nominalization.

Even though we agree that Duanmu’s Nonhead stress rule captures the general patterns of the syllabic restrictions on nominalizations – the Y position, being an NP is mapped onto a phonological phrase (\( \varphi \)) (see Selkirk 2002 among others), it over-generates in that all the constituents that occur in the post-de position must be disyllabic. As we mentioned in the last section (in examples (10b) – (10d)), monosyllabic nouns are perfectly permitted in the same position. Only words that can be independently used as verbs (or “derived nouns” following Fu 1994 and Fu et al. 2001) are subject to the syllabic restrictions. The Nonhead stress rule fails to distinguish the so-called “derived nouns” from “regular nouns”.

Moreover, we take issue with Anti-allomorphy in (20) for at least three reasons: (a) it requires that the lexicon registers pairs of lexical items which are (near) synonymous; and (b) as we have seen in (8), contrary to what Duanmu (2007) and Deng (2021) claim, nominalized verbs do not necessarily have synonymous parts; and (c) Anti-allomorphy yields the wrong predictions, as we demonstrate below.

The prediction made by Anti-allomorphy when a monosyllabic verb does not have a disyllabic counterpart is that the monosyllabic verb will be able to be nominalized. The example Duanmu (2007) provides to support this is sǐ ‘die/be dead’ in (21), which according to him does not have a disyllabic synonymous counterpart. In such a case, Anti-allomorphy does not apply, thus allowing sǐ to appear in the nominalized slot (i.e., the nonhead position), with sǐ considered to have an empty beat (S∅).

(21) tā de sǐ
3sg de die
‘his death’

This line of reasoning predicts that all monosyllabic words that do not have disyllabic counterparts should be able to be nominalized because the Anti-allomorphy Rule is not violated. In fact, it does not apply at all. In fact, sǐ ‘die’ in (21) is among the few exceptions in Mandarin that monosyllabic words are nominalized. The other two are àì ‘love’ and hèn ‘hate’ (see Lu 1981). Moreover, as shown in (22), a common verb such as chī ‘eat’ cannot be nominalized even though it does not have a disyllabic counterpart.

(22) a. Tā chī le jī.
3sg eat prf chicken
‘He ate some chicken.’

b. *duì jī de chī
prf chicken de eat
\textit{Intended: ‘eating chicken’}

c. *tā de chī
3sg de eat
\textit{Intended: ‘his eating’}

A more subtle but nonetheless important counterexample concerns chá ‘to investigate’, which we have seen above as a monosyllabic counterpart of diàochá ‘to investigate’. We have mentioned above that chá can also mean ‘to check, examine’ as in chá zìdiǎn ‘to check the dictionary’. This meaning of chá has no disyllabic counterpart. Nonetheless, as we can see in (23), it is not possible to nominalize chá under this reading.

\textsuperscript{11}The disyllabic sìwáng is typically translated as ‘death, doom’. It seems to be more nominal than verbal. Further, sìwáng is typically cast as a comparison with ‘existence’ or ‘survival’.

\textsuperscript{12}It seems to be the case that the monosyllabic verbs that can be nominalized are all statives. See also footnote \textsuperscript{7} concerning adjectives, which are also statives.
To wit, the previous phonological/prosodic accounts either overgenerate such that all the post-*de* constituents should be disyllabic (such as Fu’s comparison to English -*er* and Duanmu’s Non-head stress rule), or incorrectly predict that monosyllabic verbs without a disyllabic synonym can be nominalized (such as the Anti-allomorphy rule). The next subsection will connects the disyllabic restriction to the complexity of word structures.

### 3.2. Disyllabic restriction and structural complexity

The discussion so far concerning the syllabic restriction centers upon post-*de* nominalization. In fact, the same disyllabic restriction applies to the post-classifier constituents, as shown in (24).

Crucially, only disyllabic words can be used irrespective of the types of classifiers. *Cì* in (24a) is a typical “verbal” classifier which measures events, which Fu (1994) calls “process classifiers”. In contrast, *piān* in (24b) is a nominal classifier which measures entities. Following Grimshaw (1990), Fu (1994), and Deng (2021), *diàochá* ‘investigate/investigation’ in (24a) should be taken as a process event noun (i.e., the process of investigation) while that in (24b) denotes the result of the process (i.e., the investigation report).

(24) a. nà cì diàochá/*chá  
DEM CLF investigate  
‘that investigation’

b. nà piān diàochá/*chá  
DEM CLF investigate  
‘that investigation (report)’

Moreover, the same syllabic restriction can be observed when the “event verbs” follow auxiliaries such as *jìnxiāng* ’proceed’ (see also the end of section 2.2), they are obligatorily disyllabic, as shown in (25).

(25) duì nà-ge ànjìān jìnxiāng diàochá/*chá  
PREP DEM-CLF case undergo investigate  
‘investigate the case’

This indicates that the disyllabic restriction is not constrained to post-*de* nominalization contexts. Rather, for verbs to appear in typical nominal contexts, the disyllabic restriction is observed. Assuming the Exo-Skeletal framework (Borer 2005a,b, 2013, 2014), we argue below that the disyllabic restriction can be reduced to the restriction of word structure complexity in terms of re-categorization: disyllabic nominalized verbs in Mandarin always involve a derived complex structure. Recall Fu’s two lines of inquiry that we mentioned in the introduction, the current proposal explores her second line of inquiry with an updated structural analysis.

Consider first the basic structure starting with *de*, which we assume to select a nominal element (see also Cheng and Sybesma 2009). Following the Exo-Skeletal framework, it is interpreted as *de* defines its complement as an N. The *de*-phrase in a simple example such as ([10b]) would be as in (26).
Next, it should be pointed out that most of the examples that we have seen so far, as well as most of the nominalization examples in the literature in Mandarin involve on the surface zero-derivation. That is, there is no overt nominalizing counterpart such as -ation in English. We follow Borer (2013), who argues against zero-derivation in English, in exploring an analysis of Mandarin nominalization which avoids zero-derivation. This issue will be revisited at the end of the paper.

To understand the reasoning behind our analysis, consider first cases of nominalization which arguably do not involve zero-derivation, as in (27). All these nominals are headed by 方法 ‘method, way’.

(27) some nominals headed by 方法 ‘method, way’
    看-方法 see-method ‘opinion’
    管-方法 manage-method ‘means’
    加-方法 add-method ‘addition’
    写-方法 write-method ‘style of writing’

Assuming that 方法 is a nominal element that turns a verbal element into a noun, we suggest that 看-方法 ‘opinion’ (i.e., way of seeing) has the following representation:

![Diagram](https://example.com/diagram.png)

In other words, we treat 方法, as a C-functor which shifts the category from V to N, similar to -ation as in formation in English. We propose to extend this to cover nominalized structures which seemingly do not involve a nominal functor or a nominal head. Accordingly, the structure of 调查 ‘investigate/investigation’ following de would be as in (29).

---

11We do not think that these cases are derived by root-merge. An elaborated argumentation against root-merge is beyond the scope of the current paper, keen readers are referred to Cheng and Cheng (2021) for a brief discussion.
The immediate “difficulty” of extending the analysis illustrated in (28) to disyllabic nominalized verbs such as diàochá ‘investigate’ is that there is no obvious “nominalizer” similar to fǎ. As is clear in (29), chá is therefore treated as a C\textsuperscript{N[V]} in the same manner as we treat fǎ in (28). Considering the fact that diàochá is also a verb, one might object to the treatment of chá as a C\textsuperscript{N[V]}. Nonetheless, we would like to point out some similarities between nominalization with fǎ ‘method’ and other disyllabic nominalizations without a clear “nominal” head, as well as between the structure in (29) and the verbal structure of diàochá ‘investigate’. Consider first the examples in (30) and (31), repeated from (8b).

(30) Some disyllabic words with chá ‘investigate’
- diào-chá investigate\textsuperscript{\textcircled{14}}-investigate ‘opinion’
- kǎo-chá test-investigate ‘inspect, investigate’
- xún-chá patrol-investigate ‘go around to investigate’
- fù-chá duplicate-investigate ‘investigate again’

(31) Some disyllabic words with lùn ‘discuss’
- tǎo-lùn discuss-discuss ‘to discuss’
- biàn-lùn debate-discuss ‘to debate’
- zhēng-lùn argue-discuss ‘to argue’
- píng-lùn evaluate-discuss ‘to comment’
- tán-lùn talk-discuss ‘to talk about’

(30) and (31) illustrate that there are many combinations based on the second element in a disyllabic word/verb, similar to the case of fǎ. We have noted above that many cases of the nominalized verbs involve a modification relation between the first element and the second element (e.g., lùn in (31)). Such modification relation in the case of verbal structure means that we are dealing with modification of events, which in structural terms, involves root adjunction. Following Marantz (2013) and others (see also the discussion of denominal verbs below), the way to represent the verbal lùn-examples in (31) is as in (32) (taking as a short cut, a D\textsuperscript{(ir)}M(orphology) type of representation).

\textsuperscript{14}It should be noted that though it is often said that diào means ‘investigate’, it has other meanings as well, such as ‘transfer, move, shift’.
Returning to the issue of the disyllabic restriction, based on the analysis of disyllabic nominalized verbs such as diàochá ‘investigate’, we can now offer a structural account. First of all, why does nominalization require disyllabic words? On the basis of the structure illustrated in (29), de selects a C\_N[V] and this C\_N[V] must select for a Verb-equivalent element. Recall that according to Exo-Skeletal framework, C-functors cannot be null. As each morpheme corresponds to a syllable in Mandarin, a complex structure like (29) naturally requires a disyllabic combination. The same reasoning applies to post-classifier (see (24)) and post jìnxíng ‘proceed’ (see (25)) contexts, both of which also take N-equivalents as their complements.

Further, why do we only observe the syllabic restriction on “nominalized verbs” but not regular nouns? This is the very problem that any phonological account cannot satisfactorily resolve. As mentioned above, de, as well as classifiers, typically selects for nominal complements. We propose that de and classifiers either directly define an N-equivalent root or take a constituent that has been defined as an N-equivalent by a C-functor (i.e., C\_N[V]). When a root is defined, it is not subject to the syllabic restriction because no extra functor is needed. Hence, a monosyllabic word is permitted, as illustrated in (33) (cf. (29)). Again, due to the existence of category-changing functor, words are obligatorily disyllabic in nominalization cases.

(33) a. tā de shū
   3SG de book
   ‘his book’

   b. deP
      de C=N
      [C\_N[V] √shū]
      ‘book’

3.3. A loose end: Internal VP structure in nominalization

Fu (1994), whose work is set prior to the Exo-skeletal model put forth by Borer, argues for an internal verb phrase inside the nominalization structure with a list of diagnostics. For instance, as shown in (34), a duration phrase can occur inside the VP.
(34)  a.  Zhāngsān duì zhè-ge ànjìàn (de) liǎng-tiān de diàochá jǐnxíng de hěn shùnlì.
    Zhangsan PREP this-CLF case DE two-day DE investigation proceed DEG very smoothly.

   'Zhangsan's investigation of this case for two days went smoothly.'

b.  N’
    VP  N
    NP  V’  V_i  ∅
    PP  V’

   t_i  duration phrase

However, consider non-derived nominals such as tàidù ‘attitude’. Although it cannot co-occur with a duration phrase (see (35)), it can indeed pass the other VP-tests that Fu (1994) uses. These facts re-open the discussion concerning whether de-nominalization really involve a verbal structure.

(35)  *liǎng-tiān de tàidù two-day DE attitude

We have indicated above that in Chinese languages, everything in a noun phrase precedes the noun, with de playing a central role. First, consider the object argument of the verb. In the process nominals that we have seen above and also in Fu (1994), the object always shows up with duì, which is either a preposition or an applicative head (see Cheng and Sybesma 2015 and Tsai 2016). It turns out that it is not difficult to find examples with a duì-phrase in the pre-de position in the case of non-derived nominals, as in (36a) (cf. (36b) in the case of de-nominalization).

(36)  a.  duì zhè-jiàn shì de tàidù
    PREP DEM-CLF matter DE attitude
    ‘the attitude towards this matter’ duì-phrase

b.  duì zhèi-jiàn shì de diàochá
    PREP DEM-CLF matter DE investigate
    ‘the investigation of this matter’

Fu (1994) indicates that de-nominalization also allows temporal phrases, and subjects. (37a) and (38a) show that both temporal phrases and subject-like phrases can appear preceding de in the case of non-derived nominals (cf. the b-examples with derived nominals).

(37)  a.  qùnián de zhèngcè
    last.year DE policy
    ‘last year’s policy’ temporal phrase

b.  qùnián de diàochá
    last.year DE investigate
    ‘last year’s investigation’
Consider now examples which put the pre-de elements together.

(39) a. Zhāngsān duì wàirén de tàidū bù hǎo.
   ‘The attitude that Zhangsan has towards outsiders is not good.’

   b. dānwèi qùnián duì yuánɡōnɡ de tàidū bù hǎo.
   ‘The attitude of the unit towards the workers was not good last year.’

(39a,b) show that even for a non-derived nominal such as tàidū ‘attitude’, it is quite easy to pile up various (argument-like) constituents to the left of de. Lastly, consider the test that Fu (1994) uses to distinguish between a result reading and a process reading: by using the classifier cì ‘occurrence’ (see also the discussion in section 2.2). The idea is that zhè-cì diàochá [this-time investigation] differs from zhè-piān diàochá [this-CLF investigation] in that the latter refers to the result of the investigation (e.g., a report) whereas the former concerns the process of investigation. Nonetheless, it is possible to use cì with tàidū ‘attitude’, too, as in (40a,b).

(40) a. Tā zhè-cì de tàidū bù hǎo.
   ‘His attitude this time is not good.’

   b. zhè-cì tā de tàidū bù hǎo.
   ‘His attitude this time is not good.’

As we have indicated above, the demonstrative-classifier sequence can have intervening modifiers, as in (41), which is similar to the noun phrase in (40b).

(41) nà-běn [Zhāngsān de] shū
   ‘that book of Zhangsan’s’

In short, Mandarin nominalization under discussion (e.g., diàochá ‘investigate/investigation’) may have a comparable internal structure with non-derived nominals such as tàidū ‘attitude’. In other words, they may not necessarily involve a full VP structure as proposed by Fu (1994). Though crucial for understanding the internal structure of Mandarin nominalization, we will leave it as a loose end for further research since whether there is a full VP structure does not undermine the current analysis of the syllabic restriction on nominalization.

4. Denominal verbs

We now turn to the question of why diànhuà ‘telephone’ cannot form denominalized verb while diàn can. We have mentioned above that though verbs are heads, they do not have to be monosyllabic. In fact, Duanmu (2007:185) shows that although 73% of “old” verbs in Mandarin vocabulary are monosyllabic, only 2% of the newly created verbs are monosyllabic. He suggests that both nouns and verbs in the new vocabulary are mostly disyllabic, irrespective of being heads or non-heads in structural terms. This is indeed contrary to what the metrical analysis would predict. In other words, a pure phonological or prosodic factor cannot rule out
disyllabic denominal verbs. Instead, we suggest that the restriction rests upon a combination of (structural) factors.

Consider first what we have indicated above concerning the disyllabic verbs. We noted that one of the ways to make disyllabic verbs involves a modifying root. We extend this to disyllabic nouns such as diànhuà ‘telephone’, in that diàn ‘electric’ serves a modifying function to huà ‘talk’; see (42) for other elements that serve to modify huà ‘talk’.

\[(42)\]
\[
a. \ fǎn-huà \\
   \text{reverse-talk} \\
   \text{‘ironic remark’} \\
\]
\[
b. \ dui-huà \\
   \text{opposite-talk} \\
   \text{‘dialogue’} \\
\]
\[
c. \ xián-huà \\
   \text{leisure-talk} \\
   \text{‘gossip’} \\
\]

If the first element of the nouns in (42) is indeed a modifying root, diànhuà ‘telephone’ has a structure such as (43), where diàn modifies the n.\(\text{\textsuperscript{15}}\)

\[(43)\]
\[
\text{\textsuperscript{n}} \rightarrow \text{\textit{diàn}} \rightarrow \text{\textsuperscript{n}} \rightarrow \text{\textsuperscript{huà}}
\]

Denominal verbs, under an analysis such as Harley (2005), have a structure such as (44), where the categorizer \(v\) (i.e., \(v_{\text{cat}}\)) takes a root (projected to \(\sqrt{P}\)).

\[(44)\]
\[
\text{\textit{vP}} \rightarrow \text{\textsuperscript{DP}} \rightarrow \text{\textsuperscript{v'}} \rightarrow \text{\textit{v}} \rightarrow \text{\textsuperscript{\sqrt{P}}} \rightarrow \text{\textsuperscript{\sqrt{drool}}}
\]

Assuming that diànhuà ‘telephone’ is already categorized as an \(n(P)\), \(v_{\text{cat}}\) would not be able to take diànhuà as a root, but rather as an \(nP\), yielding the structure in (45), with \(n\) projecting to \(nP\). Since diànhuà ‘telephone’ is already categorized, we hypothesize that it cannot be categorized again by \(v_{\text{cat}}\). Instead, we can insert a “light verb” dã ‘hit’ to form dã-diànhuà ‘to telephone’ (46).

\(\text{\textsuperscript{15}}\)We are using DM representations in this section so that it is easier to see its connection with discussions of denominal verbs in the literature. The corresponding structures that we would propose following the Exo-Skeletal framework is similar to \[\text{\textsuperscript{25}}\].
This raises the question of whether there is any N to V type of operation, which can be considered to be category-changing operation. We think that noun-incorporation is in fact such an operation (see Baker 1988 among others). In English, for example, babysit, grocery-shop are considered to be possible examples of noun-incorporation though in English this is not a productive process. In Chinese languages, there does not seem to be any syntactic noun-incorporation. We think that cases such as chī-fàn [eat-rice] and zǒu-lù [walk-road] involve pseudo-incorporation (see Dayal 2011 and Borik and Gehrke 2015). In other words, Chinese languages choose to insert a “light” verb in cases such as (45) instead of incorporating the noun.

We have seen above that Chinese languages have very restricted types of denominals; in particular, mainly instrumental nouns can also be verbs. Following Marantz (2013) among others, we assume that instrumental roots such as √hammer in English directly adjoins to v. We extend this to cover Mandarin instrumental denominals: instrumental roots can also adjoin to v, as illustrated with shū ‘comb’ in (47).

Note that even if we can interpret diàn-huà ‘telephone’ as an instrument (as a communication instrument), it cannot be adjoined to v, because it is not a root; instead, it consists of a derived structure (i.e., already categorized roots), as illustrated in (48).

---

This correctly predicts the ungrammaticality of diànhuà in (15), repeated here as (49).

(49) Yǒu shénme wèntí, qǐng diàn/*diànhuà fúwúchù.
    have what question please phone reception
    ‘If you have any question, please call the reception.’

Consider now diàn in (49). If we extend the instrumental root analysis by adjoining √diàn to v as in (50), we can derive the grammatical monosyllabic version in (49). It should further be noted that diàn itself can be a denominal verb, with the meaning to give or get an electric shock. In the context of (49), √diàn should be taken as the truncated form of diànhuà, carrying the interpretation of the disyllabic word diànhuà, namely ‘telephone’. As noted in Duanmu (2007:164), that disyllabic words can be shortened to monosyllables through truncation is a common practice in Mandarin which also contributes in great measure to the word-length flexibility.

(50) vP
    v
    √diàn vcat fúwúchù

It should also be noted that in the structure in (50), due to the fact that the noun phrase diànhuà is not selected by vcat, the noun phrase fúwúchù ‘reception’ can appear in the complement position of vcat (cf. the structure of dǎ-diànhuà ‘to telephone’ in (46)), yielding the correct prediction for (49).

As another illustration, we discuss below the monosyllabic denominal verb, bāo ‘bag, to pack’, as illustrated in (51).

(51) a. bāo sān-běn shū
    bag three-CLF book
    ‘to pack three books’

b. *dā-bāo sān-běn shū
    hit-bag three-CLF book
    ‘to pack three books’

c. bǎ zhè sān-běn shū dā-bāo
    BA DEM three-CLF book hit-bag
    ‘to pack the three books’

Consider first the contrast between (51a) and (51b). Even though dā-bāo and bāo have similar meanings, only the latter can directly take an object. Taking √bāo as an instrumental root, (51a) has the structure as in (52), with √bāo ‘bag’ directly adjoining to v.

(52) vP
    v
    √bāo vcat sān-běn shū

In contrast, in the case of dā-bāo, we would have the following structure, where the complement position of v is already occupied by the instrumental NP.
The contrast between bāo and dǎ-bāo lends support to our analysis that there is no movement from the instrumental root to vcat. That is, instrumental roots have to adjoin to v in order to be verbalized.

Before closing off this section, it is worth pointing out that disyllabic denominal verbs are not completely impossible in contemporary colloquial Mandarin. For instance, Bǎidù ‘baidu’ and Gǔgē ‘google’ can directly function as verbs in (54). As sheer transliteration, both words are mono-morphemic. Hence, they can be taken as roots which can accordingly adjoin to v.

(54) Yǒu shénme wèntí jiù Bǎidù/Gǔgē yī-xià.
    ‘If you have any question, just search it on Baidu/Google.’

In this section, we have shown that the monosyllabic restriction on denominal verbs is closely related to the fact that instrumental roots are allowed to adjoin to v. Since disyllabic elements are generally no longer roots, they cannot be adjoined to v. We leave open the question of why roots other than instruments (e.g., the counterparts of drool, sweat) cannot be selected by vcat (or be verbalized by it).

5. Conclusion

We have presented above structural accounts of the disyllabic restrictions on nominalization and denominal verbs in Mandarin. In particular, using the Exo-Skeletal framework, we treat the second element of the nominalized verbs as a kind of nominalizer (i.e., C-functor in this framework), avoiding an empty categorizer. Before we conclude, we would like to revisit an updated structure à la Fu (1994) under the DM framework. In section 3.3, we have reproduced Fu’s structure in (34b). In that structure, the verb moves rightward to adjoin to the N, which is null. In (55), we update this structure under the DM framework, where the verb diàochá undergoes leftward movement to adjoin to the empty noun, and the vP eventually moves up to the left of de (presumably to the Spec deP).

(55)
It should be noted that the verb *diàochá* ‘investigate’ is composed by adjoining the first root directly to the verb *chá*, just as we suggested for these combinations. This whole verbal constituent then moves to adjoin to *n*.

Consider now how this structure differs from the structure that we proposed above in (29). In (55), the nominalizer is null and the verb moves to adjoin to it. In principle, it does not matter whether the verb is monosyllabic or disyllabic, it should be able to undergo this movement; and we have already discussed earlier that verbs can be mono- or di-syllabic. In other words, based on this structure, it is not clear why there would be a disyllabic restriction.

One might suggest that we treat *chá* on a par with our analysis in (29). Under the DM framework, as far as we can see, we would have to insert √*chá* into *n*, as in (56). This is of course a completely legitimate step, as we can insert other roots into *n* even without the *vP* (e.g., √*huā* ‘flower’).

(56)

Nonetheless, this analysis raises a couple of issues. First, it implies that *chá* as a root can be inserted in *n*, making it in fact impossible to rule out a monosyllabic “verb” following *de*. Second, even though *diào* seems to be similar to *chá* in meaning, it nonetheless cannot select DPs such as ‘this case’ because *diào zhè-ge ànjìànn* is simply ungrammatical in Mandarin. It implies that the structure (56) makes the wrong prediction in selection.

In our account under the Exo-Skeletal framework, the reason why the disyllabic restriction can be explained is due to the fact that a $C_{N[V]}$ requires a verbal complement. This ensures that an extra element has to be present. That is, the complexity of the structure corresponds to complexity of the word. In contrast, within the DM, unless we start adding notations to *n*, it is impossible to distinguish an *n* which categorizes, and an *n* which strictly requires that a *vP* follows. Crucially, insertion of null categorizers (e.g., *n* or *v*) will not set any restriction on the complexity of word structures, and thus the syllable numbers of words.

Our analyses of both nominalization and denominals still leave a lot of open questions. In the case of nominalization, future work needs to consider other V-V combinations, and explore the functor $C_{N[V]}$. In the case of denominals, it is important to understand why only instrumentals are productive. We hope that this work can lead to further discussions on this topic.
References


Borik, Olga, and Berit Gehrke. 2015. An introduction to the syntax and semantics of pseudo-incorporation. In *The syntax and semantics of pseudo-incorporation*, 1–43. BRILL.


