Mandarin Bare Indefinites

Jianan Liu, Xiaoli Dong, Bert Le Bruyn

UiL-OTS

The ability of bare nouns (BNs) to take indefinite readings has recently come to the fore again in relation to Russian (Borik 2019; Šimik & Demian 2020). We complement this literature by zooming in on Mandarin, where bare nouns have been argued to take indefinite readings with some restrictions. These restrictions have been captured at the level of syntax in terms of an S/O asymmetry (e.g. Cheng & Sybesma 1999) and at the level of discourse in terms of referential opacity (e.g. Wright & Givón 1987). We focus on the hitherto unnoticed restriction exemplified in (1a/b). *Diànyǐng* can take an indefinite reading whereas *tǎlóu* cannot. Crucially though, both are in the object position of the same verb and their (in)definiteness can be assessed at the sentence level. This contrast can thus not be explained by appealing to an S/O asymmetry or to referential opacity.

(1) a. *Wǒ   zài   kàn   **diànyǐng***   b. *Wǒ zài   kàn   **tǎlóu.***
      I   ASP see   movie   I   ASP see   tower
      'I am watching **a movie**.'   'I am looking at **the tower**.'

We aim to capture the contrast in (1a/b) while preserving the insights from Cheng & Sybesma (1999) and Wright & Givón (1987). With the former, we agree that indefinite readings of BNs are limited to the object position. From the latter, we take over the insight that there is a fine-grained competition between BNs and nouns marked by the numeral *yi* ('one'). Rather than analyzing this competition at the discourse level, we do so at the level of semantics.

Why we don't pursue a full-fledged pseudo-incorporation analysis

Wright & Givón's notion of referential opacity is reminiscent of that of discourse opacity in the literature on pseudo-incorporation. (2a/b) show that Mandarin BNs on their indefinite readings are more discourse opaque than on their definite readings. Whereas *diànyǐng* only allows for anaphoric pickup by a zero anaphor, *tǎlóu* prefers to be picked by a full anaphor. This contrast shows that discourse opacity has indeed a role to play in the analysis of indefinite readings of Mandarin BNs.

(2) a. *Wǒ   zài   kàn   **diànyǐng**. (#**Tā**) shì   qùnián shàngyìng de.*
      I   ASP watch movie   it be   last   year   release   DE
      'I am watching a movie$_i$. It$_i$ was released last year.'
   b. *Wǒ   zài   kàn   **tǎlóu**. #(**Tā**)   shì   táng   cháo   jiànzào de.*
      I   ASP see   tower   it   be   Tang   dynasty build DE
      'I am looking at the tower$_i$. It$_i$ was built in the Tang dynasty.'

We argue that contrasts like the one in (3a/b) warrant an $<e,t>$ analysis of Mandarin BNs on their indefinite readings but don't suffice to argue in favour of a full-fledged pseudo-incorporation analysis. To make a real case for pseudo-incorporation, we would need to find other signs that something special is going on at the syntax-semantics interface. The literature has shown that these signs can be very diverse and range from special syntactic positions for pseudo-incorporated nouns (e.g. Hungarian | Farkas & de Swart 2003) over article drop in article languages (e.g. Spanish | Espinal & McNally 2011) to lack of case marking (e.g. Hindi | Dayal 2011). None of these signs can be tracked in Mandarin as the language has rigid word order, counts as an articleless language and does not mark case overtly. This should not be taken to mean that we're not dealing with pseudo-incorporation but that we only have evidence to support a more general $<e,t>$ analysis.

Our proposal: a fine-grained relational semantics of VN combinations
The gist of our proposal is that the availability of indefinite readings of Mandarin bare nouns depends on the semantics of both the verb and the object noun. Crucially, the object noun has to come with a relational argument that the verb can pick up on and link its subject to. A simpler case than the one in (1b) is given in (3). *Dìdì* is a relational noun and comes with a relational argument. *HAVE* has been argued to pick up on relational arguments in its object noun and link its subject to it, allowing the subject to end up in the relational argument position of the object noun. Analyses of *HAVE* in English have to work around the indefinite article (e.g. Partee 1999). We argue that Mandarin is more transparent and restricts the indefinite readings of bare nouns to exactly those environments where the semantics of the verb and noun conspire to link the verb's subject to a relational argument in the nominal part of its object. (4) gives a compositional sketch.

(3) *Zhāngsān yǒu **dìdì** jiào Xiǎomíng.*　　(4)　$[[\text{HAVE}]] = \lambda R \lambda y \exists x(R(y)(x))$
　　Zhangsan have brother name Xiaoming　　　　$[[\text{BROTHER\_OF}]] = \lambda y \lambda x(\text{brother\_of}(y)(x))$
　　'Zhangsan has **a brother** named Xiaoming'　　$[[\text{Z. HAVE BROTHER\_OF}]] = \exists x(\text{brother\_of}(Z.)(x))$

Following Le Bruyn et al. (2016), we extend the analysis in (4) to more subtle cases where lexical verbs pick up on implicit arguments in their object noun's Qualia Structure (Pustejovsky 1995). In the case of *diànyǐng* in (1a), the verb *kàn* can be interpreted as picking up on an implicit argument in the telic role of the noun *MOVIE* (watch(x,y)). A similar link between an implicit argument of *TOWER* and *kàn* is not available in (1b) and this explains why *tǎlóu* only gets a definite reading.

In (1a/b), we kept the verb constant. The contrast between (5a/b) shows that our analysis also applies for minimal pairs in which we keep the noun constant. *DRESS* comes with the same telic role in (5a) and (5b) (wear(x,y)) but only *chuān* and not *ná* can arguably pick up on this role. Our analysis then predicts that *qúnzi* in (5a) but not in (5b) can get an indefinite reading. As the translations suggest, this prediction is borne out.

(5) a. *Wǒ chuān-zhe lànsè **qúnzi**.*　　　　b. *Wǒ ná -zhe lànsè **qúnzi**.*
　　I　wear ASP blue dress　　　　　　　　I　hold_up ASP blue dress
　　'I'm wearing a blue dress.'　　　　　　'I'm holding up the blue dress.'

Corpus data support: indefinites in Harry Potter in English and Mandarin
To broaden the empirical basis of our analysis, we conducted a corpus study into the role of VN combinations in the interpretation of Mandarin BNs. Rather than relying on a monolingual corpus, we used a parallel corpus based on the first Harry Potter volume and its Mandarin translation. The parallel corpus buys us an independent criterion for (in)definiteness, allowing us to restrict our attention to those expressions that correspond to *a(n) ... N* in English. We further restricted the data to those translations of *a(n) ... N* that appear in object position in Mandarin.

Our corpus contains 153 datapoints with 32 BNs and 121 nouns marked by *yi*. We annotated for the availability of (implicit) (relational) arguments in the object nouns and the ability of the verb to pick up on these. If a VN combination satisfied these requirements, we marked it as +R, if it didn't, we marked it as -R. Annotation was done by two annotators (Cohen's κ=.93). Results show that 81% of bare nouns (n=26) appeared in +R VN combinations whereas nouns marked by the numeral *yi* only appeared in +R VN combinations in 24% of the cases (n=32). These results provide strong support for the role of +R VN combinations in the availability of indefinite readings of BNs. In the talk, we present further statistical support and discuss further relevant factors.

Conclusion | The Mandarin data we analyze reveal an important role for (i) the lexical semantics of nouns and (ii) verbs as well as a highly frequent use of the (iii) numeral *yi* in the Mandarin

reference system. All three factors have largely been ignored in the analysis of reference but offer promising new perspectives on the way reference is organized across articleless languages.

Selected references | Cheng & Sybesma (1999), *LI*, 30(4). | Le Bruyn, de Swart & Zwarts (2016), *Lingua*, 182 | Wright & Givón (1987), *Studies in Language*, 11(1).