“Noun-less” nominal expressions in Mandarin Chinese

Nick Huang (znhuang@nus.edu.sg). National University of Singapore / University of Connecticut
West Coast Conference in Formal Linguistics 2020

1 Introduction

This talk: nominal syntax and classifier semantics, with a Mandarin Chinese case study.

(1) liang *(ben) shu
    two CL book
    ‘two books’ [Mandarin Chinese, Num Cl Noun]

Case study: exceptions to the above generalization.

(2) liang {nian / tian / piao / ye / bu / hua / bi}
    two year day vote page step stroke stroke
    ‘two years / days / votes / pages / steps / strokes of the brush/pen’


- Sidenote: will refer to (1) as “tripartite expressions.”

I review existing analyses of bipartite expressions and add novel arguments to make the claim that they are syntactically bipartite:

- No phonologically-null noun, NP ellipsis, etc.

If this “nounlessness” argument goes through, there are interesting implications for syntax and semantics.

Syntax:

- Data pose a problem for the NP Hypothesis of nominal expressions (Bruening 2009; Bruening et al. 2018, also Chomsky 1970).
- Indirect evidence in favor of the DP Hypothesis (Abney 1987; Szabolcsi 1994, etc.).

Semantics: Cross-linguistic distribution of classifiers has been tied to deeper cross-linguistic variation in noun and numeral semantics (Chiaretti 1998a; Rothstein 2010; Krifka 1995; Wilhelm 2008; Bale & Coon 2014; Sudo 2016b).

- Core idea: classifiers are needed so that numerals and nouns can combine with each other to express cardinality.
- Presuppose presence of numerals and nouns.
- Data pose problems for these proposals about numeral and noun semantics.
- Discuss whether we can “rescue” these proposals.
Caveat: Many other aspects of the syntax and semantics of classifiers, whether in Chinese or in other languages, that I will not be able to address here for scope and time reasons – please see the references.

2 Bipartite nominal expressions

2.1 Hypotheses

There are a number of hypotheses one could have about the representation of bipartite nominal expressions and lexical items like nian ‘year’ or hua ‘(brush)stroke’; some of the simpler ones are set out in (3). To avoid pre-judging the issue, I will henceforth refer to these lexical items as post-numeral lexical items (henceforth, ‘PNLs’).

I will argue in favour of the first hypothesis (3a): a PNL is a classifier (C. J. Tang 2005, S.-W. Tang 2012, 2013) and there is no null NP following it (cf. S.-W. Tang 2012, 2013, Kayne 2007).

(3) a. Num Cl b. Num NP + null Cl c. Num NP + head movement
d. Num Cl + elided Npe. Num Cl + covert NP

2.2 PNLs are morphosyntactically classifiers

First, reduplication (4) classifiers can be reduplicated to give a ‘one-by-one’ reading (following S.-W. Tang 2012, 2013).

(4) Zhe yi xie baogao, ta [yi fen fen] de kan.

PNLs can also be reduplicated to produce the same reading (5).


b. Shijian jiu zheyang [yi tian tian] de guoqu.

c. Women yao [yi piao piao] de qu zhengqu xuanmin we must one vote vote ADV go fight.for voter MOD support

d. Ta [yi ye ye] de fan-kan manhua shu.

e. Guojia zheng [yi bu bu] de zou xiang miewang.

Second, complementary distribution with “plural classifier” xie, which indicates a plurality of items denoted by the following NP (e.g. Li 1999).

(6) a. yi ben shu

‘one book’
b. yi xie shu / *yi ben xie shu / *yi xie ben shu  
   one PL book  one CL  PL book  one PL  CL  book  
   ‘some books’

PNLs are also in complementary distribution with plural xie (7).

(7) *yi xie (tian / hua / ye / bu)  
   one PL day stroke page step  
   ‘some days / strokes / pages / steps’

Third, a “no bare classifiers” test, after an observation by Lü 1990 [1944]  

A slight detour: three generalizations about Mandarin Chinese:

- The numeral yi ‘one’ can be omitted (although interpreted) (8a).
- NPs can be omitted through ellipsis (8b).
- However, the omission of yi ‘one’ and an NP cannot happen together,  
  i.e. classifiers cannot appear alone (8c).

(8) a. Ta mai-le (yi) ben shu.  
   3SG bought one book  
   ‘He bought one book.’

b. Ta mai-le yi ben shu, wo ye mai-le yi ben (shu).  
   3SG bought one CL book I also bought one CL book  
   ‘He bought one book, and I also bought one (book).’

c. Ta mai-le ben shu, wo ye mai-le ben *(shu).  
   3SG bought CL book I also bought CL book  
   ‘He bought one book, and I also bought one book.’ (Yang 2001:76, ex. 33)

Like classifiers, PNLs cannot appear alone (9).

(9) a. Zhe ge zi shao-le *(yi) hua.  
   this CL character missed one stroke  
   ‘This character is missing a stroke.’ (e.g. pointing out mistakes in Chinese characters.)

b. Wo zai zhe lüguan zhu-le *(yi) tian.  
   I at this hotel stayed one day  
   ‘I stayed at this hotel for a day.’

c. Lisi huode *(yi) piao.  
   Lisi receive one vote  
   ‘Lisi received a vote.’

d. Zhe fen biaoge lou-le *(yi) ye.  
   this CL form missed one page  
   ‘This form is missing a page.’

e. Wo tui-le *(yi) bu.  
   I moved one step  
   ‘I took a step back.’

2.3 PNLs are not nouns

(3) c. Num NP + head movement of NP to Cl (cf. Simpson & Ngo 2018 for Vietnamese)

\[
\begin{align*}
\text{Num} & \quad \downarrow \quad CL+N \\
& \quad \downarrow \quad \text{NP} \\
& \quad \downarrow \quad nian/\text{hua}
\end{align*}
\]

First, as a rule, nouns can co-occur with the plural xie (6). But (7) showed  
that PNLs cannot.

(6) yi xie shu  
   one PL book  
   ‘some books’

(7) *yi xie (tian / hua / ye / bu)  
   one PL day stroke page step  
   ‘some days / strokes / pages / steps’

Second, nouns can appear bare in Mandarin Chinese (10). But PNLs cannot,  
as we saw in the discussion of the “no bare classifiers” test (9).
Third, nouns (and even overt pronouns) can be modified with relative clauses (11).

(11) a. \([RC \text{ hui youyong de}] \text{ ren} / \text{ wo}
   \text{ can swim MOD person I}
   ‘a person who can swim’ / ‘I, who can swim’

b. \([RC \text{ xie cuo de}] \text{ bihua}
   \text{ write wrong MOD brushstroke}
   ‘a stroke that was written incorrectly’

PNLs cannot be modified directly by relative clauses (12).
(To modify a PNL, one must first quantify it with a numeral to produce a nominal expression (12b)).

(12) a. \(*[RC \text{ tamen yin cuo de}] ye
   \text{ they print wrong MOD page}
   \text{ Intended: ‘a page that they printed incorrectly’}

b. \([RC \text{ tamen yin cuo de}] yi ye
   \text{ they print wrong MOD one page}
   ‘a page that they printed incorrectly’

2.4 There is no null NP in bipartite expressions

2.4.1 Ruling out NP ellipsis

(3) d. Num Cl + elided NP

For concreteness, suppose the hypothesis in (3d) works as follows:
- There is a set of covert (phonologically null) nouns that mean ‘time,’ ‘vote,’ ‘(brush) stroke,’ etc.
- PNLs are classifiers; the form of the PNL depends on the noun.
- E.g., \text{hua} is used with the null noun that means ‘stroke.’

One might expect covert nouns to have the same properties as phonologically overt nouns. This prediction is too strong.
First, overt nouns can be modified with relative clauses. But these covert nouns (if they exist) cannot be, even though the PNL makes clear which noun is intended.

NP ellipsis is usually unacceptable in out of the blue contexts.

(13) \text{Biang zi} \text{ you 42 ge [NP #bihua].}
\text{biang character has 42 CL brushstroke}
‘The character biang has 42 brushstrokes.’

Bipartite expressions are acceptable in out of the blue contexts.

(14) \text{Biang zi} \text{ you 42 hua.}
\text{biang character has 42 stroke}
‘The character biang has 42 brushstrokes.’

2.4.2 No covert noun


\begin{center}
\begin{tikzpicture}
  \node (num) {Num};
  \node (cl) [below left of=num] {CL};
  \node (np) [below right of=num] {NP};
  \node (nianhua) [below of=cl, xshift=-2cm] {nian/hua};
  \node (zero) [below of=np, xshift=2cm] {\_\_\_};
  \draw (num) -- (cl) -- (np);
  \draw (num) -- (nianhua);
\end{tikzpicture}
\end{center}

For concreteness, suppose the hypothesis in (3d) works as follows:

- There is a set of covert (phonologically null) nouns that mean ‘time,’ ‘vote,’ ‘(brush) stroke,’ etc.
- PNLs are classifiers; the form of the PNL depends on the noun.
- E.g., \text{hua} is used with the null noun that means ‘stroke.’

This assumption about null nouns is already somewhat problematic, as phonologically null elements are usually posited for closed-class items or functional heads, and less often for open-class morphemes like nouns.
(15) a. (i) liang ge [\textit{RC} xie-cuo de] \textit{bihua} two CL write-wrong MOD brushstroke ‘two brushstrokes that were incorrectly written’
   (ii) *liang hua [\textit{RC} xie-cuo de] \(\theta\textit{stroke}\) two CL\(\textit{stroke}\) write-wrong MOD
b. (i) shi zhang [\textit{RC} pingshen tou de] \textit{xuapiao} ten CL judge cast MOD ballot ‘ten votes that judges (in a contest) cast’
   (ii) *shi piao [\textit{RC} pingshen tou de] \(\theta\textit{vote}\) ten CL\(\textit{vote}\) judge cast MOD

c. (i) [\textit{RC} tamen yin-cuo de] \(\theta\textit{page}\) yi ye they print-wrong MOD one CL\(\textit{page}\) ‘a page that they printed incorrectly’
   (ii) *yi ye [\textit{RC} tamen yin-cuo de] \(\theta\textit{page}\) one CL\(\textit{page}\) they print-wrong MOD

Second, we would predict that the plural “classifier” \textit{xie} can appear with covert nouns. This is not the case (16), even there is a suitable context.

(16) a. *Xiaohai xie-cuo-le yi \textit{xie} \(\theta\textit{stroke}\) child write-wrong-PFV one PL
   Intended: ‘The child wrote some strokes incorrectly.’ (Context: observing a child who is just starting to learn Chinese characters.)
   b. *Zhe ge cansai zhe jingran hai neng yingde yi \textit{xie} this CL participant unexpectedly still can \textit{xie} win one PL
   \(\theta\textit{vote}\)
   Intended: ‘This participant still somehow managed to win some votes.’ (Context: describing someone who performed poorly in a contest and should have not won any votes from the judges)

Third, an argument from ellipsis.

- Generalization: ellipsis requires some degree of semantic identity be-

- If so, covert nouns should license ellipsis of overt nouns that are semantically identical (17b)
- But ellipsis is not possible here.

(17) a. Zhe ge \textit{zi} shao-le liang \textit{ge bihua}, na \textit{ge} this CL character missed two CL brushstroke that CL \textit{zi} ze shao-le yi \textit{ge bihua}. character while missed one CL brushstroke ‘This character is missing two strokes, while that character is missing one (stroke).’
   (The first \textit{bihua} licenses ellipsis of the second \textit{bihua})
   b. #Zhe ge \textit{zi} shao-le liang \textit{hua} \(\theta\textit{brushstroke}\), na \textit{ge} this CL character missed two \(\theta\textit{brushstroke}\) stroke that CL \textit{zi} ze shao-le yi \textit{ge bihua}. character while missed one CL brushstroke (Null \(\theta\textit{brushstroke}\) does not license ellipsis of \textit{bihua})

(18) a. Zhe ge houxuanren yigong huode 4010 \textit{zhang xuapiao}, this CL candidate a.total.of win 4010 CL ballot duishou \textit{ze} zhi huode 105 \textit{zhang xuapiao}. rival while only win 105 CL ballot ‘This candidate won a total of 4010 votes, while his/her rival only won 105 (votes).’
   b. #Zhe ge houxuanren yigong huode 4010 \textit{piao} \(\theta\textit{vote}\), this CL candidate a.total.of win 4010 vote duishou \textit{ze} zhi huode 105 \textit{zhang xuapiao}. rival while only win 105 CL ballot

\(^3\)There are arguments that semantic identity is insufficient. For instance, Merchant (2013) argues that structural parallelism is also necessary. But it is not clear that structural parallelism is the issue here, unless we intend it to include lexical identity.
2.5 **Interim summary**

- No clear evidence that there is NP ellipsis or a noun.
- Most parsimonious explanation: bipartite expressions are nounless syntactically.

3 **Consequences**

3.1 **Syntax of nominal expressions**

Recent work by Bruening 2009; Bruening et al. 2018 (reviving Chomsky 1970) argue that nominal expressions are NPs – headed by nouns.

- Bruening et al. (2018) apply this analysis to classifier languages like Korean and Vietnamese (19).
- This might be the case for Korean and Vietnamese (but see Simpson & Ngo 2018), but difficult to extend to noun-less bipartite expressions in Chinese.

(19) NP

\[
\begin{array}{c}
D \\
\text{ClP} \\
\text{Num} \\
N \\
\text{Cl}
\end{array}
\]

By Occam’s Razor, the data indirectly favour the DP Hypothesis / functional head hypothesis over the NP Hypothesis.

3.2 **On the semantics of classifiers**

Debate on the semantics of classifiers, nouns, and numerals.

- Numerals need to pair up with nouns to express cardinality.
- In certain languages, numerals or nouns have some semantic “deficiency” that prevents them from pairing up directly.
- Classifiers are needed to remedy that deficiency.
- Dispute over whether the deficiency lies in numerals or nouns (Bale & Coon 2014).
One perspective, “classifiers for nouns” (e.g. Chierchia 1998a,b, see also Rothstein 2010).

- Numerals can only combine with count nouns.
- Mandarin only has mass nouns (some equivalent thereof), while English has count and mass nouns.
- A classifier ‘individuates [mass nouns] to a level suitable to counting.’ (Chierchia 1998a, p. 93)
- In formal terms, following Bale & Coon 2014, pp. 696–697:

(21) English \[\text{\texttt{two}}\] = Mandarin \[\text{\texttt{liang}}\] = \(\lambda P \lambda x. \text{ATOM}(P). \{x :*P(x) & \mu_{\theta}(x) = 2\}\).

(22) a. \[\text{\texttt{brushstroke}}\] = \(\{x : \text{ATOM}(x) & \text{BRUSHSTROKE}(x)\}\)
   b. \[\text{\texttt{bihua}}\] = \(\sqcap \text{BRUSHSTROKE} \) (the “brushstroke kind”)

(23) Classifier \[\text{\texttt{ge}}\] = \(\cup \) (a function from kinds to sets of atoms)

Another perspective: “classifiers for numerals”: variation in the semantics of numerals.

- All languages use a measure function to express cardinality.
- English numerals have this measure function “built in” (Križka 1995).
- Mandarin numerals do not, and so require a classifier to supply it.
- Adapting Wilhelm 2008, p. 55:

(24) a. \[\text{\texttt{two}}\] = \(\lambda P \lambda x. [P(x) & \text{OU}(x) = 2]\)
   OU = function that gives the cardinality of ‘object units’
   b. \[\text{\texttt{liang}}\] = 2

(25) \[\text{\texttt{brushstroke}}\] = \[\text{\texttt{bihua}}\] = \(\lambda x. [\text{BRUSHSTROKE}(x)]\)

(26) \[\text{\texttt{ge}}\] = \(\lambda n \lambda P \lambda x. [P(x) & \text{OU}_{\text{general}}(x) = n]\)

Both proposals presuppose the presence of numerals and nouns — current data pose a problem.

One conclusion one might draw is that the cross-linguistic distribution of classifiers cannot be boiled down to cross-linguistic differences in noun semantics (or numeral semantics).

- Maybe there is nothing deep here. Languages just vary in whether their nominal expressions require classifiers.

3.3 Salvaging the hypotheses?

I would like to suggest another way of thinking about the hypotheses. And if we do that, there might be an argument in favor of the “classifiers for numerals” hypothesis.

The two hypotheses can be re-framed as claims about classifier semantics: A central part of the “classifiers for numerals” proposal is that the denotation of classifiers contains a measure function. Classifiers take numerals and nouns as arguments (26).

So why is there no noun in these ‘nounless’ expressions?[4]

- Hypothesis: the classifiers here (PNLs) have a rich enough semantics in their own right, so no need for a noun argument.
  - For example, \texttt{hua}'s denotation contains the measure function and a noun-like semantics that approximately means ‘brushstroke.’

In contrast, the intuition behind the “classifiers for nouns” hypothesis: numerals occur with a nominal if it has count noun semantics.

---

[4] There are also instances where classifiers do not appear with an overt numeral in Mandarin and other varieties of Chinese, as Bale & Coon (2014), Zhang (2019), Cheng & Sybesma (1999), Lü (1990 [1944]) observe. But one could argue on reasonable grounds that there is a null numeral (Zhang 2019) or null (in)definiteness operator that the classifier takes as an argument (Bale & Coon 2014).

In addition, under the “classifiers for numerals” hypothesis, classifiers are usually assumed to form a constituent with numerals before merging with nouns. In the context of Mandarin, this would mean that nominals are left-branching. However, I note that the hypothesis, as adapted in this section, does not make as strong a commitment.
Classifiers are best thought of as operators so that a predicate can get count noun semantics (such as Rothstein 2010; Bale & Coon 2014, implicitly in Chierchia 1998a).

So, to account for the ‘nounless’ data:

- Option 1: All classifiers, including PNLs, are operators. PNLs appear with covert mass nouns (or something equivalent).
  - Difficult to reconcile with the data above.

- Option 2:
  - Classifiers fall into two distinct semantic classes: most are operators; the rest (PNLs) have count noun semantics.
  - Disjunctive: logically coherent, but arguably not aesthetically ideal.

If this argument goes through, Mandarin would constitute another language that supports the ‘classifiers for numerals’ hypothesis (Déné Suliné, Wilhelm 2008; Mi’gmaq and Chol, Bale & Coon 2014; Japanese, Sudo 2016a).

4 Conclusion

- Reviewed properties of Chinese ‘nounless’ nominal expressions.
- Argued that these expressions do not contain a noun in the syntax: the lexical item that follows the numeral is a classifier.
- Showed that they pose a problem for recent proposals that nominal expressions are headed by N.
- Showed that they pose a problem for certain theories about the cross-linguistic distribution of classifiers, but argued that the problem might not be fatal for a ‘classifiers for numerals’ proposal.

References


Lü, Shuxiang. 1990 [1944]. "Ge-zi de yingyong fanwei, fulun danweici qian yi-zi de tuoluo [the uses of ge and omission of yi before classifiers]. In Lü Shuxiang Wenji [Collected works of Lü Shuxiang], volume 2, 144–175. Beijing: Commercial Press.


5 Appendix

5.1 PNLs with the classifier ge

Some time-interval PNLs can appear optionally with the classifier ge, e.g. xiaoshi / zhongtou ‘hour’, xingqi / libai ‘week.’ Some speakers report similar alternations for tian ‘day’ and nian ‘year.’

(27) liang (ge) xiaoshi
    two CL hour
    ‘two hours’
It is possible that these terms are syntactically ambiguous between noun and a classifier; when analyzed as a noun, the classifier ge appears (following C.-C. J. Tang 2005).

5.2 Distinguishing PNLs from classifier omission

Mandarin allows classifiers to be omitted in formal registers (e.g. news headlines), producing Num Noun sequences.

(28) \[
\text{Num Liang} [\text{Noun du-fan} \text{ jiechi [Num wu} [\text{Noun renzhi}]
\]
two drug-addict hold five hostage
‘Two drug addicts hold five people as hostages.’

But PNLs are acceptable in non-formal registers, and so are not instances of classifier omission.

- Rules out an analysis like (3b).

(3) b. \[
\text{Num} \text{ CL} / 0 \text{ NP nian/hua}
\]

5.3 Relative clauses, xie, and covert nouns

(15) and (16) show us that relative clauses and xie do not co-occur with covert nouns (if they exist).

Can we explain these examples by claiming that relative clauses and xie have some idiosyncratic requirement for an overt head noun?

- Not quite: relative clauses and xie can co-occur with elided NPs.

(29) a. Lisi mai-le yi xie shu, wo ye mai-le yi xie shu.
Lisi bought one PL book I also bought one PL book
‘Lisi bought some books, and I also bought some (books).’

b. \[
[RC Ni chang de] ge bu bi [RC wo chang de] you sing MOD song not than I sing MOD ge hao ting.
song good hear
‘The song you sang is not as nice as the one I sang.’

10