Determiners, Bare Nouns, and Donkey Sentences in Numeral Classifier Languages

Peter Jenks
jenks@berkeley.edu

Some unresolved questions:

- How is definiteness encoded in languages which lack definite articles, for example, in numeral classifier languages? (Chen 2004, Simpson et al. 2011)
- Are putatively definite nominals in classifier languages DPs or NPs? (Chierchia 1998, Cheng and Sybesma 1999, Wu and Bodomo 2009)

Claims in this paper:

- Definiteness associated with bare nouns vs. determiners in classifier languages is semantically distinct, licensed by either uniqueness or familiarity, respectively (cf. Schwarz 2009).
- The uniqueness associated with bare nominal definiteness follows from the lexical semantics of nouns in classifier languages, meaning determiners are not needed.
- The familiarity associated with DP definiteness demonstrates that what is common to determiners in across languages is the introduction of a discourse index or variable for binding, not argument creation (pace Longobardi 2005).
- These claims straightforwardly predict the distribution of these expressions in donkey sentences in numeral classifier languages.

Structure of the talk
Act I: Bare nouns in classifier languages
Act II: Two types of definites in classifier languages
Act III: Covarying interpretations of noun phrases

1 Background: Bare Nouns in Numeral Classifier Languages

Three readings of bare nouns in Mandarin Chinese and Thai:

1. Kind-based readings (1)
2. Scopeless indefinite readings (2)
3. Definite readings (3)
(1) Kind-level readings
   a. [shìzǐ hěn kuài jiù huì juézhǒng]
      lion very quick then will be extinct
      ‘Lions will be extinct very soon.’
      (Mandarin, Cheng and Sybesma 2012, ex. (6a))
   b. [mǔ] klài  sǔnphan
      mice nearly extinct
      ‘Mice are nearly extinct.’
      (Thai, Piriyawiboon 2010, p. 42)

(2) Scopeless indefinite readings
   a. Yuehan zài zhāo [yīshēng]
      John PROG looking-for doctor
      i. ‘John is looking for a(ny) doctor.’
      ii. *‘There’s a specific doctor that John is looking.’
      iii. ‘John is looking for the doctor.’
      (Mandarin, cf. Yang 2000, p. 26)
   b. chán kamlaŋ hāa [mǎa] yùu
      1SG PROG looking-for doctor IMPF
      i. ‘I’m looking for a doctor.’
      ii. *‘There’s a specific doctor that I’m looking for.’
      iii. ‘I’m looking for the doctor.’
      (Thai)

(3) Definite readings
      dog want cross road
      ‘The dog wants to cross the road.’
      (Cheng & Sybesma 1999:510)
   b. [mǎa] kamlaŋ hàw.
      dog PROG bark
      ‘The dog(s) are barking.’
      (Piriyawiboon 2011:43)

However, specific indefinite readings emerge with numerals/indefinite articles:

(4) Specific indefinite readings
   a. Yuehan zài zhāo [yì-ge yīshēng]
      John PROG looking-for one-CLF doctor
      i. ‘John is looking for a(ny) doctor.’
      ii. ‘There’s a specific doctor that John is looking for.’
      (Mandarin)
   b. chán kamlaŋ hāa [mǎo khon niŋ] yùu
      1SG PROG looking-for doctor CLF INDEF IMPF
      i. ‘I’m looking for a doctor.’
      ii. ‘There’s a specific doctor that I’m looking for.’
      (Thai)

> The ‘true’ indefinites in (4) resist the kind level readings in (1).

> Interpretive differences between bare nouns and indefinites have been argued to follow from analyses of bare nouns as bare NPs (Carlson 1977, Chierchia 1998, Yang 2000, Piriyawiboon 2010, Jenks 2011).
Additionally, the claim that demonstratives are not definite articles in classifier languages comes from the lack of consistency effects (Löbner 1985):

(5)  
\[ \text{d`ek} \ \text{khon nán ṇon ĭuù} \ \text{têo d`ek} \ \text{khon nán mai.dài ṇon ĭuù} \]
\[ \text{child CLF that sleep IMPF but child CLF that NEG sleep IMPF} \]
\[ \text{‘That child is sleeping but that child is not sleeping.’ (cf. #the)} \]  
(Piriyawiboon 2011:49)

Thus, Jiang (2012:15) claims that “most [classifier languages] do not have marked definiteness…”

In this light, two views of definite bare nouns:

1. Definite bare nouns are NPs; definiteness arises somehow via semantics (e.g. Chierchia 1998, Yang 2000, Piriyawiboon 2010)
2. Definite bare nouns are full DPs, or some equivalent (Cheng and Sybesma 1999, Visonyanggoon 2000, Simpson 2005, Wu and Bodomo 2009)

View 2 assumes that bare definites are semantically equivalent to definite descriptions in, e.g., English.

> Yet definite bare nouns are only available for some readings that are available for definite descriptions.

## 2 Two Types of Definites in Numeral Classifier Languages

Two competing views of definiteness in the semantics literature:

**Uniqueness** Definite descriptions are used for referents that are unique (e.g. Russell 1905):

(6)  
\[ \text{a. The sun is hot.} \]
\[ \text{b. The woman that John went out to dinner with last night was brilliant.} \]

**Familiarity** Definite descriptions are used for referents that are familiar, or are anaphoric to referents that have been mentioned earlier (e.g. Heim 1982):

(7) John bought a book and a magazine. The magazine was expensive.

Recent cross-linguistic work has established that both notions might be needed in different languages:

- Schwarz (2009): ‘Strong’ vs. ‘weak’ definite articles in German dialects licensed by familiarity or uniqueness, respectively.
- Arkoh and Matthewson (2013): Describe the Fante determiner nU, licensed only by familiarity.
- Schwarz (2013): Reviews a number of other languages where a similar distinction holds.

The same kind of generalization applied to numeral classifier languages:

(8)  
<table>
<thead>
<tr>
<th>Category</th>
<th>Definition</th>
<th>Licensed by</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare definites</td>
<td>Bare nouns with definite translations in English</td>
<td>Uniqueness</td>
</tr>
<tr>
<td>Indexical definites</td>
<td>Pronouns, demonstrative descriptions, complex definites</td>
<td>Familiarity</td>
</tr>
</tbody>
</table>
2.1 Bare definites must be unique

Four environments license bare definites (I will illustrate these just with Thai, but in all the cases I have checked, they hold for Mandarin as well):

Environment 1 Globally unique definites:

(9) \[\text{moon (CLF that) bright very}\]

\[\text{‘The moon is very bright.’}\]

(Thai)

Likewise, ‘unique kinds’ are translated with bare nouns: (NB: Thai provinces get one Senator, two MPs)

(10) a. \[\text{M.P. Chiang Mai CLF that angry very}\]

\[\text{‘#The/That MP from Chiang Mai is very angry.’}\]

(Thai)

b. \[\text{senator Chiang Mai CLF that angry very}\]

\[\text{‘The/#That Senator from Chiang Mai is very angry.’}\]

(Thai)

Environment 2 Weak definites (Poesio 1994, Carlson et al. 2006):

(11) a. \[\text{Suthep took Somchai to the hospital}\]

\[\text{‘Suthep took Somchai to the hospital.’}\]

(Thai)

b. \[\text{Suthep took Somchai to the building}\]

\[\text{‘Suthep took Somchai to the building.’}\]

(Thai)

> (11b) requires a context to be natural, but not (11a).

> Schwarz (2014) provides an analysis of weak definites which relies on uniqueness relevant to an event.

Environment 3 Bridging/associative definites:

(12) \[\text{If PROSP go clinic POSS doctor S. recommend COMP should PROSP ascend stairs}\]

\[\text{‘If your going to Dr. Somchai’s clinic, I recommend that you take the stairs.’}\]

(Thai)

(13) \[\text{That car was intercepted by the police because there wasn’t a sticker on the license plate.’ (Thai)}\]

> These bridging contexts can be taken to be situationally unique.
Environment 4 High animate referents; these are also exceptional, apparently allowing bare definite uses even when familiar:

(14)  a. \( S1: \) fiaan sàmpaat phɔ̀o lûuk khûu niŋ.
    H. interview father son CLF-pair INDEF
    ‘Hans interviewed a father-son pair.’

    b. \( S2: \) kháw khît wàa \( (\text{tua})-phɔ̀o (\#\text{khon nán}) \) mây nàa-s̀ôncay
    He think COMP SELF-father CLF that NEG interesting
    ‘He didn’t think the father was interesting.’

(15)  a. \( S1: \) fiaan sàmpaat nàk-khîn khon niŋ \( \text{kap nàk-kaanmìaŋ khon niŋ.} \)
    H. interview writer CLF INDEF and politician CLF INDEF
    ‘Hans interviewed a writer and a politician.’

    b. \( S2: \) kháw khît wàa \( (\text{nàk-kaanmìaŋ} (\#\text{khon nán}) \) mây nàa-s̀ôncay
    He think COMP politician CLF that NEG interesting
    ‘He didn’t think the politician was interesting.’

> Basic human terms like ‘man’ or ‘woman’ see optional use of the demonstrative; these have formed the basic examples in the literature, but they are exceptional relative to other kinds of nouns.

> Following Zwarts (2014), these ‘high animacy’ definites may be seen as instantiating unique roles in functional frames (in fact, this generalization would subsume the earlier two cases as well).

SUMMARY: Bare nouns are restricted to uniqueness definites once uniqueness is relativized to situations/events/frames.

2.2 Indexical definites must be anaphoric

Indexical definites must be anaphoric or strongly familiar in the sense of Roberts (2003).

(16)  a. \( S1: \) zuótiān wò yù-dào \( \text{yī ge xuéshèng} \)
    Yesterday 1SG meet one-CLF student
    ‘Yesterday I met a student.’

    b. \( S2i: \) \( \text{nà ge xuéshèng} / \text{tā} \) hên còngmìng.
    that CLF student / 3SG very clever
    ‘That student(s)he was very clever.’

    c. \#S2ii: \( \text{xuéshèng} \) hên còngmìng.
    student clever very
    ‘Students are clever.’

(Mandarin)
Indexical definites are discourse-anaphoric rather than directly referential:

(18) De dicto definites (cf. ‘Hob-Nob sentences,’ e.g. Elbourne 2005: p. 6)

   man kill Suthep.  
   ‘A man killed Suthep.’

b. S2i: tamruat sōngsāy wāa [phūu-chaay khon nān]/[khāw] dāy-rāp bāat-cēp tōn-nān  
   police suspect man CLF that / 3SG receive injury time-that  
   ‘Police suspect that that man was injured at the time.’

   police suspect man receive injury time-that  
   ‘Police suspect that a man was injured at the time.’ (Thai)

Indexical definites must pick out the reference of the antecedent, rather than its sense:


a. S1: pii-nūi naayōk pen samāachik phūa-Thai.  
   year-this Prime Minister pred member pro-Thai.  
   ‘This year the prime minister is a pro-Thai party member.

   year-next 3sg / P. M. CLF that FUT pred member democrat  
   ‘Next year he will be a democratic party member.’

   year-next Prime Minister FUT pred member democrat  
   ‘Next year he will be a democratic party member.’ (Thai)

A puzzle: Why are indexical definites ‘strongly anaphoric’ in Thai/Mandarin?

2.3 Analysis: Uniqueness versus familiarity

(20) a. Bare definites = Bare, kind-denoting NPs, saturated with a situation variable. 

b. Indexical definites = Full, presuppositional DPs which introduce an index.
Bare Definites

- Bare nouns are kinds, functions from situations to individuals (cf. Krifka 1995, Chierchia 1998).
- Definiteness arises by saturating the kind with the default ‘topic situation’ (Schwarz 2009):

\[(21) \ \ \ \ \ \begin{array}{l}
\text{a. NP} \\
\quad \text{NP} \\
\quad \quad \text{nàa} \\
\quad \quad \quad \text{‘dog’} \\
\text{b. } [nàa]^\theta = \lambda s. \text{DOG}(s) \\
\text{c. ‘the total set of dog(s) in some situation’}
\end{array}\]

- Situation variables are freely adjoined to major syntactic constituents.
- Situation Restriction returns the largest set of dogs in a particular situation (the supremum) (Jiang 2012, a.o.).
- Weak and uniqueness uses of definites follow directly provided some adequate theory of situations Kratzer (cf. 1989), Elbourne (cf. 2013).

Indexical Definites

- Demonstratives are choice functions \(f_c\) (Reinhart 1998).
- Demonstratives introduce an additional argument which corresponds to a discourse markers (dynamic indices = \(u_i\)) (Groenendijk and Stokhof 1990, Chierchia 1995, Schwarz 2009).
- Demonstratives presuppose existence, but not uniqueness (explaining (5)).

\[(22) \ \ \ \ \ \begin{array}{l}
\text{a. DP} \\
\quad \text{DP} \\
\quad \quad \text{DP}_4 \\
\quad \quad \quad \text{DP}_4 \\
\quad \quad \quad \quad \text{DP} \\
\quad \quad \quad \quad \quad \text{u}_4 \\
\quad \quad \quad \quad \quad \quad \text{Clf} \\
\quad \quad \quad \quad \quad \quad \quad \text{nán} \\
\quad \quad \quad \quad \quad \quad \quad \quad \text{khon} \\
\quad \quad \quad \quad \quad \quad \quad \quad \text{person} \\
\quad \quad \quad \quad \quad \quad \text{nákrian} \\
\quad \quad \quad \quad \quad \quad \quad \quad \text{‘student’} \\
\quad \quad \quad \quad \quad \quad \quad \quad \text{Clf} \\
\quad \quad \quad \quad \quad \quad \quad \quad \text{Clf} \\
\quad \quad \quad \quad \quad \quad \quad \quad \text{StUDENT} \\
\text{b. } [[\text{DP}_4]^\theta = \lambda s.f_c x : \exists x \forall \text{STUDENT}(x)(s)[\forall \text{STUDENT}(x)(s) \land x = g(u_4)] \\
\text{c. ‘that student who is discourse referent 4 in some situation’}
\end{array}\]

- In actual deictic/pointing uses of indexical definites the discourse marker is identified with a real-world object.
- The reliance on an index explains the ‘rigid’ properties of indexical definites: The variable assignment is held constant across situations.
- Paycheck readings for English definites are available because uniqueness licenses the in English but not indexical definites in classifier languages; this effect bleeds over into pronominal reference.
3 Covarying interpretations of noun phrases

(23) Two more generalizations:
   a. Indexical definites freely receive covarying (bound) interpretations.
   b. Bare nouns cannot receive covarying interpretations

(24) Donkey anaphora in Mandarin (loosely based on Cheng and Huang 1996)
   a. ruguo ni kandaof [yi ge xuesheng], qing jiao na ge xuesheng/la (ta3) lai jian wo.
      if you see a CLF student please tell that CLF student 3SG come see me
      ‘If you see a student, please tell that student to come see me.’
   b. ruguo ni kandaof [yi ge xuesheng], qing jiao xuesheng+1/gen lai jian wo.
      if you see a CLF student please tell student come see me
      ‘If you see a student, please tell students to come see me.’

(25) Donkey anaphora in Thai
   a. [chaawnaa thuk khon thi mii khwaay] tii [khwaay tua nan] / man /
      farmer every CLF that have buffalo hit buffalo CLF that / it
      ‘Every farmer that has a buffalo hits it.’
   b. [chaawnaa thuk khon thi mii khwaay] tii [khwaay+1/gen]
      farmer every CLF that have buffalo hit buffalo
      ‘Every rice farmer that has a buffalo hits buffalo.’ (a generic claim)

The same effects arise in the absence of a unique referent for the donkey anaphor:

(26) [baan thuk laa thi mii doy] jaa mii doy ik baj tag yuu khay-khay mii /
      house every CLF REL have water.basin PROSP have water.basin ADD CLF set IPFV next.to EC /
      ?man / doy baj nan / *doy
      it water.basin CLF that water.basin
      ‘Every house that has a water basin has another water basin sitting next to it.’

Two analyses of donkey anaphora:

Situation Binding  Donkey anaphora are not bound like regular variables, but receive covarying interpretations via bound situation variables (Elbourne 2005, 2013)

Dynamic Binding  Donkey anaphora are interpreted as regular bound variables, and mechanisms allow indefinites to escape their typical scopal restrictions (e.g. Groenendijk and Stokhof 1990, Chierchia 1995).

In this section I show:

1. Bare nouns are problematic for the situational approach.
2. No apparent problems for the dynamic approach.

→ The crucial difference between these theories is that in dynamic approaches, but not in ‘pure’ situation-based approaches, donkey anaphora are instances of normal variable binding.

→ My basic claim, then, is that DPs, but never NPs, supply a variable for binding.

1 Again, there are systematic exceptions to this generalization involving bridging contexts which refer to part-whole relationships as well as high animacy individuals. Cf. Schwarz (2009) for similar facts once again in German article systems.
3.1 Problems for the situational approach to donkey anaphora

Elbourne (2001, 2005): Pronouns are concealed definite DPs; definite DPs are bindable under all the same configurations as pronouns:

(27) Every man who owns a donkey beats the donkey.

Elbourne (2013): Covarying interpretations of definites arise purely through situation semantics:

(28) a. Every man who owns a donkey beats the donkey.

b. ‘In every situation $s_1$ where a man owns a donkey, there is another situation $s_2$, where $s_2$ is a subpart of $s_1$ and in $s_2$ a man beats the unique donkey in $s_1$.’

> Because $s_1$ is restricted to donkey-owning situations, only the donkeys in $s_1$ can be beaten.

But if the bare noun $khwaay$ ‘buffalo’ in (29) denotes the unique student or set of students in a particular situation, Elbourne’s analysis incorrectly predicts it should allow bound interpretations:

(29) [chaawnaa thuk khon thi mii khwaay] tii [khwaay*1/gen] farmer every CLF that have buffalo hit buffalo

‘Every rice farmer that has a buffalo hits buffalo.’ (a generic claim) (Thai)

3.2 Dynamic approaches to donkey anaphora

G&S 1990: “The essential feature of [Dynamic Predicate Logic] is that it allows an existential quantifier to bind occurrences of variables outside its scope…”

(30) Every man who owns a donkey beats it.

a. $\forall x[\text{man}(x) \land \exists y[\text{donkey}(y) \land \text{owns}(x,y)]] \rightarrow \text{beat}(x,y)$

b. $\forall x\forall y[[\text{man}(x) \land \text{donkey}(y) \land \text{owns}(x,y)] \rightarrow \text{beat}(x,y)]$ (DRT-derived meaning)

c. $\forall x[[\text{man}(x) \land \exists y[\text{donkey}(y) \land \text{owns}(x,y)]] \rightarrow \exists y[\text{donkey}(y) \land \text{owns}(x,y) \land \text{beat}(x,y)]]$

(Chierchia 1995, p. 124)

The dynamic approach makes the right prediction: because donkey anaphora are normal bound variables, indices should be necessary for binding.

(31) [chaawnaa thuk khon thi mii khwaay] tii [khwaay tua nán] / [man] farmer every CLF that have buffalo hit buffalo CLF that

‘Every farmer that has a buffalo hits it.’ (Thai)

(32) a. $[\text{khwaay tua nán}] = \text{THAT} x[\neg \text{BUFFALO}(x) \land x = y]$ (simplified from (22b))

b. $[\text{(34)}] = \forall x[[\text{farmer}(x) \land \exists y[\text{buffalo}(y) \land \text{own}(x,y)]] \rightarrow \exists y[\text{THAT} z[\neg \text{BUFFALO}(z) \land z = y] \land \text{own}(x,y) \land \text{beat}(x,y)]]$

c. “For all $x$ such that $x$ is a farmer and there is some $y$ such that $y$ is a buffalo and $x$ owns $y$, then there exists some $y$ such that $y$ is identical to some buffalo $z$ and $x$ beats $y$.”

Bare nouns cannot be bound because they lack the variable characteristic of indexical definites.
4 Conclusions

(33) Implications for theories of definiteness and determiners:
   a. Familiarity and uniqueness are both necessary components of a theory of definiteness, confirming Schwarz (2009).
   b. The UG-encoded contribution of determiners in is of an index for variable binding rather than argument formation.
   c. A Prediction: Uniqueness definites are expressible without determiners, but anaphoric definites may never be.

(34) Implications for analyses of numeral classifier languages:
   a. The distribution of definite interpretations in classifier languages conforms to cross-linguistic patterns of definite marking.
   b. Bare nominals are different from English definite descriptions; their semantics follows from a kind-based meaning. Hence, there is no compelling semantic reason to think that definite bare nouns are DPs.
   c. Viewing indexical definites as strictly anaphoric accounts for puzzles in the ‘strict’ interpretation of pronouns in numeral classifier languages. If pronouns and demonstratives are determiners (and there are good reasons to think they are), numeral classifier languages can project DP.
   d. Numeral classifiers themselves are somewhat irrelevant to definiteness (pace Cheng and Sybesma 1999, 2005, 2012, but Cantonese and Vietnamese must be closely examined along these lines to be sure); numeral classifiers simply produce sets from kinds that can then combine with choice functions.

References

Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral Diss., University of Massachusetts, Amherst.


