Are there “weak” definites in bare classifier languages?*

Ka-Fai Yip
Yale University

Ushasi Banerjee
Yale University

Margaret Chui Yi Lee
University of Connecticut

Abstract  This paper motivates a new view on the typology of definiteness that integrates (quasi-)names. The primary data is drawn from Cantonese and Bangla, where both bare classifier constructions and bare nominals are recruited for definite expressions. We argue that these bare nominals, while often analyzed as the so-called “weak”/unique definites in other languages, are indeed name-like expressions akin to the quasi-name Mom in English, in contrast with the definite descriptions denoted by bare classifier constructions. We propose that quasi-names, as well as proper names, are derived by a definite determiner that encodes a functional relation between the discourse participants and the referent. We further discuss cases where quasi-names compete with definite descriptions and proper names. The findings not only suggest that names should be brought into the picture, but also shed light on how pragmatic principles interact and determine the choice of referring expressions.

Keywords: definiteness, quasi-names, bare nominals, the unique-anaphoric dichotomy, functional relation, competition

1 Introduction

An important part of the study of definiteness is understanding the different strategies adopted in natural languages to encode it (Schwarz 2009; Arkoh & Matthewson 2013; Jenks 2018, i.a.), and the factors that influence the choice of a particular definite

* Acknowledgment: We wish to thank the four anonymous SALT reviewers for their valuable comments. We are also very grateful to Florian Schwarz, Andrew Simpson, Diti Bhadra, Zoltán Szabó, and the members of Yale Semantics Reading Group for their help, in particular Veneeta Dayal, for her guidance and continuous support. For discussions, we thank Eno Agolli, Dorothy Ahn, Simon Charlow, Elizabeth Coppock, Jon Gajewski, Maribel Romero, Uli Sauerland, and the audience at Yue-25 (Jinan), UConn, South Asian Semantics reading group, and SALT 33 (Yale). For judgment and comments, we thank Ka-Wing Chan, Sheila Shu-Laam Chan, Tommy Tsz-Ming Lee, Carmen Kin Man Tang for Cantonese; Arko Banerjee, Kousani Banerjee, Diti Bhadra, Nirnimesh Bhattacharjee, Ishani Guha, Ankana Saha for Bangla; and Comfort Ahenkorah for Akan. We are especially thankful to the SALT editors for their help. All errors remain our own.

©2023 Yip, Banerjee, Lee
form in a given context (e.g., Ahn 2019, 2023, cf. Heim 1991). In this study, we address these two questions by investigating definite expressions in two bare classifier languages, Cantonese and Bangla, where definiteness is encoded in two ways. The first and predominant way is to use the so-called bare classifier construction (a term due to Simpson 2005, henceforth bare CL). A bare CL is a combination of a classifier and a noun, that denotes a definite description (CL-N in Cantonese, Cheng & Sybesma 1999; N-CL in Bangla Bhattacharya 1999; Dayal 2012). The second way is to employ bare nominals (henceforth bare N) in certain uniqueness contexts (Simpson, Soh & Nomoto 2011).

By probing into the distinction between bare CLs and bare Ns, we argue that these two languages offer a new perspective on the typology of definiteness as well as the competition among referring expressions.

In his seminal work, Schwarz (2009, 2013) proposes a dichotomy to capture two important notions in definiteness: uniqueness (Frege 1892; Russell 1905) vs. familiarity/anaphoricity (Heim 1982; Roberts 2003). The dichotomy is manifested as two forms of definite articles in German and Fering. The weak article vom ‘by-the-weak’ in German is used for unique referents in a situation, whereas the strong article von dem ‘by the strong’ is used for establishing anaphoric links to a referent previously mentioned in the discourse. We refer to the former as unique definites and the latter as anaphoric definites. This line of research has been pursued in a wide range of languages, where bare Ns are claimed to be unique definites, and a separate form to be an anaphoric definite, such as determiners in Akan (Arkoh & Matthewson 2013, but see Bombi 2018; Owusu 2022), bare CLs in Bangla (Biswas 2014, but see Simpson & Biswas 2016 for nuances), demonstratives in Mandarin (Jenks 2018, but see Dayal & Jiang 2022; Simpson & Wu 2022), among others. Importantly, the typology informs us about the competition between different referring expressions, as in the choice between bare Ns and determiners (Owusu 2022), demonstratives (Jenks 2018; Ahn 2019), and pronouns (Ahn 2019, 2023), with recruitment of principles like Maximize Presupposition! (Heim 1991), among others.

In this study, we argue that Cantonese and Bangla do not fit into the current typology based on the “unique-anaphoric” dichotomy. Rather, the difference between definite bare Ns and bare CLs corresponds to the contrast between (quasi-)names and definite descriptions (like the NP in English) (see also Cheng & Sybesma 1999; Jenks 2018 for a similar idea alluded to for Cantonese). Quasi-names refer

---

1 Apart from the definite reading, bare Ns can also convey kind/generic readings in both languages, which we set aside for future research.

2 The term “weak definites” in this paper, when occasionally used, refers to uniqueness-base definites, rather than Carlson’s sense of “weak definites” (Carlson, Sussman, Klein & Tanenhaus 2006). The latter are definites that have an indefinite-like reading, such as the newspaper in Lola is reading the newspaper (but see Schwarz 2014 for a possible unification with unique definites).

3 For other forms of distinction, see Schwarz (2019) and Royer (2022) and the references therein.
to name-like expressions that carry descriptive content, such as the capitalized use of *Mom* in English (Pelczar & Rainsbury 1998), or *school* as in *School is closed today.*\(^4\) We show that definite bare Ns in both languages behave like names as rigid designators. Additionally, they encode a functional relation between discourse participants and the referent (as in ‘our Mom’). Bare CLs, on the other hand, denote definite descriptions covering both unique and anaphoric definites. This motivates the need for a new typology that integrates (quasi-)names, as illustrated in Table 1. The revised typology adds a significant piece to the understanding of the choice among definite forms: names, as we demonstrate, compete with definite descriptions, along with demonstratives and pronouns. Furthermore, we bring in an understudied aspect of the topic: the competition between quasi-names and proper names.\(^5\)

<table>
<thead>
<tr>
<th>Language</th>
<th>Type</th>
<th>Definite description</th>
<th>Quasi-names</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cantonese</td>
<td>CL-lang.</td>
<td>bare CLs</td>
<td>bare Ns</td>
</tr>
<tr>
<td>Bangla</td>
<td>CL-lang.</td>
<td>bare CLs</td>
<td>bare Ns</td>
</tr>
</tbody>
</table>

**Table 1** The typology of definiteness with quasi-names (preliminary)

This paper is organized as follows. §2 shows that the unique-anaphoric dichotomy does not capture the distinction between bare Ns and bare CLs. §3 introduces the functional relation in bare Ns, which sets them apart from unique definites. §4 further probes into the name-like properties of the bare Ns, which set them apart from definite descriptions. §5 proposes a compositional quasi-name analysis of bare Ns; their denotation differs from definite descriptions at both DP and NP levels. §6 discusses the role of names in the competition between referring expressions. Finally, §7 concludes with implications on definiteness.

2 The flawed uniqueness-anaphoric dichotomy

At first glance, an apparent strong-weak pattern seems to exist in Cantonese and Bangla. In anaphoric cases, where strong articles are used in German, only bare CLs can be used to denote anaphoric referents as in ‘the principal’ in (1).

---

\(^4\) This notion of quasi-names should be distinguished from “text-internally licensed quasi-names” in Kim (2023), whose referent has a linguistic antecedent (as a subclass of anaphoric definites).

\(^5\) While it is reported that definite bare Ns in other bare CL languages also show differences with bare CLs (e.g. Simpson et al. 2011 for Hmong and Vietnamese; Simpson 2017 for Jinyun (Wu Chinese)), whether these differences are the contrast argued for here requires further careful examination.
(1) **Anaphoric:** ✔ Bare CL vs. ✗ Bare N
      Ka-Ming yesterday saw one-CL-principal and one-CL-teacher.
      {go-haauzoeng / haauzoeng} hou houjan. [C(antonese)]
      **‘KM met a principal and a teacher yesterday. The principal was very kind.’**
      Robi-r one-CL principal and one-CL teacher-GEN with see
      holo. {headmaster-Ti / headmaster} duschintay chilen.
      happen. principal-CL / principal worried AUX
      ‘Robi met a principal and a teacher. The principal looked worried.’ [B(angla)]

In contrast, in uniqueness cases, where weak articles are used in German, only bare Ns are used to refer to the unique referent in the situation, as in ‘the principal’ in (2).

(2) **Situation-uniqueness-teacher:** ✗ Bare CL vs. ✔ Bare N
   Context: A new colleague has joined the school you have been teaching at, and
   you are responsible for guiding him/her. This morning, when the new colleague
   and you arrive at the school, you tell him/her:
   a. {go-haauzoeng / haauzoeng} wui bei fan sigaanbiu nei
      CL-principal / principal will give CL timetable 2SG
      ‘The principal will give you the timetable.’
   b. {headmaster-Ti / headmaster} toma-ke nishchoi timetable-Ta diyech-en?
      principal-CL / principal you-DAT of.course timetable-CL give-PERF-3
      ‘The principal must have given you the timetable?’

While the above contrast seems to parallel the anaphoric-uniqueness dichotomy in German/Fering and may lead one to conclude that bare CLs are anaphoric definites and bare Ns are unique definites (e.g. Biswas 2014), a closer inspection shows that this is not the entire picture. Consider (3), which is also an example of situational uniqueness. Unlike (2), it is the bare CL that gets used, but not the bare N.

(3) **Situation-uniqueness-officer:** ✔ Bare CL vs. ✗ Bare N
   Context: You are an officer in Dept. of Education. You and your colleague
   have a visit to a new school today. Neither of you have met anyone from the
   school before. When you both arrive at the school, you ask your colleague:
   a. {go-haauzoeng / haauzoeng} hai naam ding neoi?
      CL-principal principal be male or female
      ‘Is the principal male or female?’
   b. {headmaster-Ti / headmaster} tomaka-e nhwoi
      principal-CL / principal you-DAT is the principal
      ‘Is this the principal?’
Are there “weak” definites in bare classifier languages?

b. ki asha korchen {headmastar-Ta/ #headmastar} kemon?
   what hope do principal-CL principal how
   ‘What are you expecting, how is the principal?’ [B]

A crucial difference exists in the two cases: in (2), the discourse participants belong to the school, and the referent is technically their principal (explicated in §3). On the other hand, in (3), the discourse participants are not affiliated with the school and the referent is not their principal. A similar difference is found in cases that involve globally unique entities. We see that ‘the moon’ in (4), which refers to the unique moon of the earth, is expressed with the bare N rather than the bare CL:

(4) Global-uniqueness-earth-moon: X Bare CL vs. ✔ Bare N
   Context: You are a parent teaching your child world knowledge. You say:
   a. {#Go-jyutloeng/ jyutloeng} hai wongfan gozan zau gin-dou.
      CL-moon/ moon at evening that.time then see-able
      ‘The moon can be seen in the evening.’ [C]
   b. {#chaand-Ta/ chaand} shondher akashe dekha dey.
      moon-CL/ moon evening sky see give
      ‘The moon appears in the evening.’ [B]

However, when ‘the moon’ is of a planet in which discourse participants do not reside, it can only be expressed with the bare CL in (5):

(5) Global-uniqueness-alien-moon: ✔ Bare CL vs. X Bare N
   Context: You are an astronaut and are performing a mission on an alien planet. There is only one moon there. You landed on the planet and collected data about the moon, and report to your team on the spaceship:
   a. {Go-jyutloeng/ #jyutloeng} hai wongfan gozan zau gin-dou.
      CL-moon/ moon at evening that.time then see-able
      ‘The moon can be seen in the evening.’ [C]
   b. {chaand-Ta/ chaand} shondher akash-e dekha dey.
      moon-CL/ moon evening sky-LOC see give
      ‘The moon appears in the evening.’ [B]

Taking stock, the distinction between bare Ns and bare CLs in Cantonese/Bangla does not correspond to the unique-anaphoric dichotomy in German/Fering, as summarized in Table 2. Instead, the choice depends on the relation between the referent and the discourse participants, which we will discuss in the next section.6

---

6 Other contrasts between unique and anaphoric definites also do not line up in Cantonese and Bangla, including immediate situation uniqueness, bridging contexts of producer-relation and of part-whole-relation (Schwarz 2009). In all these cases, bare CLs are chosen over bare Ns.
Types of definites | German/Fering Def. articles | Cantonese/Bangla Bare CL | Bare N
--- | --- | --- | ---
Anaphoric (=1) | strong | ✓ | ✗
Unique-situation-teacher (=2) | weak | ✗ | ✓
Unique-situation-officer (=3) | weak | ✓ | ✗
Unique-global-earth-moon (=4) | weak | ✗ | ✓
Unique-global-alien-moon (=5) | weak | ✓ | ✗

Table 2 The range of definites expressed by bare CLs and bare Ns

3 Functional relation

As opposed to a unique-anaphoric divide, we suggest that the difference between the use of bare Ns and bare CLs is due to a functional relation between the discourse participants and the referent. The basic intuition is informally given in (6), where a function maps the speaker and the addressee in the discourse to the referent denoted by the bare N (see Engdahl 1986; Chierchia 1992; Dayal 1996 for \( \langle e, e \rangle \) functions that map entities to entities). The formulation of this relation will be taken up in §5.

(6) \[ f_1 \text{ speaker/ addressee } \rightarrow x \ ; x \text{ is the principal in (2)} \]
\[ f_2 \text{ speaker/ addressee } \rightarrow y \ ; y \text{ is the moon in (4)} \]

To understand the nature of the functional relation that licenses bare Ns, there are a few things to note. First, it is restricted to a set of relations that apply only to individuals with the property denoted by the noun, such as principal-of-the-school-of for the bare N ‘Principal’, but not mother-of even in a context where the principal (of a different school) happens to be the discourse participants’ mother.

Second, the set of relations is not inherent in nature as in kinship relations like mother-of, since not all bare Ns are relational nouns (see §4). For instance, maaiju-lou [C] and maach-wala [B] ‘fish-seller’ may have a bare N use in a situation where a unique fish-seller exists in the relevant community. Here, the licensing relation is the-fish-seller-in-the-community-of (akin to ‘the local fish seller’). We call it a characteristic relation, as it applies only to members in the set of the property denoted by a noun, on a par with characteristic functions (to be detailed in §5).

Third, the functional relation has to hold for both speakers and addressees. In scenarios where the (characteristic) relation holds only between the speaker and the referent, bare Ns are not licensed. In (7), the referent is the speaker’s instructor but not the addressee’s, and as a consequence, the bare CL is preferred over the bare N.

...
Are there “weak” definites in bare classifier languages?

(7) Context: You have recently joined a tutorial, which your friend is not aware of. The instructor for the tutorial has announced that soon there will be a surprise quiz. Distressed, you decide to talk to that friend, who doesn’t attend the tutorial or have substantial knowledge about it. You say:

a. Nei zi-m-zi \{go-lousi/ #lousi\} gamjat gong-zo me? ...[C] 2SG know-not-know CL-teachers/ teacher today say-Pfv what ‘Do you know what the teacher said today? (There will be a quiz!)’

b. \{sir-Ta/ #sir\} ajke ki boleche janish? ... [B] teacher-CL/ teacher today what said know ‘Do you know what the teacher said today? (There will be a quiz!)’

Finally, the functional relation can be made explicit through the use of possessives. For instance, the bare Ns ‘Principal’ in (2) can be substituted with ngodei-ge haauzoeng (Cantonese)/ amader headmaster (Bangla) ‘our principal’ without changing the felicity of the sentences. The use of first-person plural possessives indicates the presence of a functional relation with the discourse participants. Likewise, the absence of such a functional relation can be indicated by the use of third-person possessives. For example, ‘the moon’ in the alien-moon case (5) can be expressed with keoidei-ge jyutloeng (Cantonese)/ oder chNaad (Bangla) ‘their moon’.

In short, we have shown that bare Ns differ from bare CLs in requiring a licensing functional relation. In the next section, we further probe into their differences and show that bare Ns are different from definite descriptions with respect to name-like properties.

4 Name-like properties

In this section, we argue that bare Ns in Cantonese and Bangla are (quasi-)names instead of regular definite descriptions. We present two arguments based on (i) scopal behavior and (ii) noun-choice restrictions as well as relation to name-marking devices. We show that bare Ns pattern with referential proper names but contrast with the definite descriptions formed by bare CLs in all these cases.

Let us consider scopal behavior first. Bare Ns, like proper names (see Muñoz 2019; Agolli 2023 and the references therein), do not take narrow scope with respect to quantificational operators. We illustrate this in three cases. The first one is a counterfactual context in (8). While the bare CL can refer to the hypothetical principal, ‘Billy’, the bare N can only refer to the elected principal in the actual world. Since the actual principal is not a murderer, using the bare N is infelicitous. That is, bare Ns cannot take narrow scope under a counterfactual conditional operator.
Counterfactual: Bare CL: ✔ vs. Bare N: ✗

Context: The principal in your school is elected by teachers. Billy lost the election last year. This year, he was found to have committed a murder. You say: “If we had voted for Billy,...”

a. ... jigaa {go-haauzoeng/ #haauzoeng} zau hai saatjanhungsau. [C] now CL-principal principal then be murderer ‘...the principal would have been a murderer.’ (#bare N: actual principal)

b. ... {headmaster-Ti/ #headmaster} ek-jon khuni hoten [B] principal CL principal one-CL murderer AUX ‘... the principal would have been a murderer.’ (#bare N: actual principal)

The second case involves universal quantification over situations, where there is a unique referent in each situation. As illustrated in (9), the bare CL can give a co-varying reading and refer to different owners/bosses who are unique in each of the restaurant-going/office-visiting situations. The reference of the bare N, on the other hand, is fixed to the unique restaurant owner or office boss in relation to the discourse participants in the actual world.

Co-variation: Bare CL: ✔ vs. Bare N: ✗

a. Ngo muici heoi caacaanteng, {go-lousai/ #lousai} dou wui 1SG every.time go restaurant CL-boss boss ALL will tung ngo kinggai. with 1SG chat ‘Every time I go to restaurants, the boss chats with me.’ [C]

b. ami jokhoni kono notun office-e jai, ontoto {boss-Ta/ #boss} 1 whenever any new office-LOC go, at least boss-CL/ boss ama-r shathe kotha bolen I-GEN with word say ‘Whenever I go to any office, the boss speaks with me.’ [B]

The third case concerns attitude predicates. In (10), the intended referent of ‘the principal’ is ‘Billy’, who the speaker thought to be a principal (i.e., de dicto), but not the actual principal ‘Mark’ (i.e., de re). While the bare CL can have the de dicto reading and refer to ‘Billy’, the bare N can only have a de re reading and refer to ‘Mark’, which is infelicitous under the given context. In other words, bare Ns cannot take narrow scope under intensional operators but necessarily denote the referent in the actual world.

Taking stock, the lack of narrow-scope reading of bare Ns in the above cases suggests that they behave as rigid designators, on a par with proper names (Kripke 1980).
Are there “weak” definites in bare classifier languages?

(10) **De dicto:** Bare CL: ✔ vs. Bare N: ✘

*Context: Maggie, a newly-appointed teacher, mistakenly identifies Billy, who she dislikes, as the principal of the school. On realizing it is in fact Mark who is the principal, she remarks:*

a. Ngo zangging gokdak {go-haauzoeng/ #haauzoeng} hai go seoijan.[C] 1SG used to think CL-principal principal be CL bad person
   ‘I used to think that the principal is a bad person.’ (true only w/ bare CL)

b. ami bhabtam [headmaster-Ti/ #headmaster] kharap manush
   I thought principal-CL/ principal bad human
   ‘I used to think that the principal is a bad person.’ (true only w/ bare CL)

We now turn to the property of bare Ns vis-à-vis a noun-choice restriction and its correlation with name-marking devices. To begin with, not all nouns have a unique bare N use. The noun choice of bare Ns is restricted to unique entities in a conventionalized context, such as ‘principal’ in a school, ‘doctor’ in a clinic, ‘church’ in a community, etc. In contrast, nouns such as ‘student’ or ‘book’ reject this use even in contexts that facilitate uniqueness like (11). Such restrictions do not apply to bare CLs.

(11) **Context:** There is only one student in the room, and the rest are teachers.

a. #(Go)-hoksaang zodai-zo. The student sat down. [C]  
   b. chhatro-#(Ta) boshlo student-CL sat  
   ‘The student sat down.’ [B]

Below, we list some nouns that typically allow a bare N use and a few that typically do not in both languages, with only the gloss given for space reasons.


Regarding the name-marking devices, Cantonese prefix *aa*- (Sio & Tang 2020) and Bangla suffix *-moshai* are typically attached to proper names, such as *aa-Gaaming ‘Ka-Ming’* in Cantonese and *Robi-moshai ‘Robi’* in Bangla. Importantly, once attached, the proper noun can no longer be used as a predicate and only have a referential reading, as in (13). Hence, the two affixes mark referential names.7

7 There are also pragmatic effects like expressing closeness to the referent for Cantonese *aa-* and respect to the referent for Bangla *-moshai.*
The set of nouns that allow a bare N use shows a striking correlation with the name-marking devices. Aa- and -moshai may combine with some common nouns to yield a name-like reading, and these nouns are the set of human nouns that have a bare N use. That is, a unique (human) bare N can always be affixed by aa-/-moshai (modulo the pragmatic constraints mentioned in footnote 7), as illustrated by ‘principal’ in (14), contrasting with ‘student’ which does not have a bare N use. Note that bare CLs never take the affixes: *aa-go-haauzoeng [C] and *headmastar-Ta-moshai [B] ‘the principal’ are ungrammatical.

(14) a. aa-{hauzoeng/*hoksaang} AA-principal/student ‘Principal/*Student’ [C] b. {headmaster/*chhatro}-moshai principal/student-HONF ‘Principal/*Student’ [B]

We suggest that aa/-moshai are proprial articles like a in Maori (Muñoz 2019), which also exclusively attaches to name NPs and necessitates a referential use (see Muñoz 2019; Agolli 2023 and references therein for the prevalence of proprial articles cross-linguistically). That bare Ns may take proprial articles like aa/-moshai indicates that bare Ns are names, specifically quasi-names along the lines of Pelczar & Rainsbury (1998) and Muñoz (2019) (e.g. Mom in ‘Mom went to the airport’). Note that quasi-names in other languages may also be attached by proprial articles, like amma ‘grandma’ or kennari ‘teacher’ in Icelandic (Sigurðsson 2006).

To sum up, we have argued that bare Ns in Bangla and Cantonese are quasi-names by showing two of their properties that are typically associated with proper names: (i) rigid designation of the referent with respect to scope, and (ii) taking name-marking devices like proprial articles. Hence, the difference between bare Ns and bare CLs corresponds to the distinction between (quasi-)names and definite descriptions.

5 Towards a quasi-name approach

In this section, we offer a compositional analysis of bare Ns as quasi-names and bare CLs as definite descriptions. We propose that their difference is manifested at two levels, DP and NP, as illustrated in (15). Combining the insights from Muñoz (2019) and Agolli (2023), we suggest that the referentiality of names (both proper and
Are there “weak” definites in bare classifier languages?

quasi-names) comes from a different definite determiner (i.e. proprial articles), and (quasi-)proper nouns themselves denote name-bearing properties (i.e. *predicativism* of names, Elbourne 2005; Matushansky 2008; Gray 2012; Fara 2015). We further propose a novel component for this determiner: apart from an *iota* operator, it also encodes a *functional relation* between the discourse participants and the referent.

(15)

<table>
<thead>
<tr>
<th>Bare CLs:</th>
<th>[ DP level</th>
<th>[ NP level   ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare Ns:</td>
<td>D₁</td>
<td>(CL⁺)Common N</td>
</tr>
<tr>
<td>Proper names:</td>
<td>D₂ with f</td>
<td>Quasi-proper N</td>
</tr>
</tbody>
</table>

![Table: Bare CLs and Bare Ns](image)

On the NP level, we suggest that there are three types of nouns: (i) common nouns (*N*ₖ); (ii) proper nouns (*N*ₚ); and (iii) quasi-proper nouns (*N*ₐ). They all denote a set of individuals as their extension and properties (relativized to a situation) as their intension, as given in (16) (*cf.* Barwise & Perry 1981; Elbourne 2005).⁸ For common nouns, the property is a descriptive content; and for (quasi-)proper nouns, the property is a name-bearing property (abbreviated as *Pₚ*) (following Elbourne 2005; Fara 2015; Agolli 2023, *i.a.*). Quasi-proper nouns additionally presuppose the descriptive content (after Muñoz 2019), and hence the term “quasi”.⁹

(16)

a. \[ [Nₖ] = \lambda s \lambda x.P(x)(s) \]
   b. \[ [Nₚ] = \lambda s \lambda x.x \text{ bears } Nₚ \text{ in } s = Pₚ(x)(s) \]
   c. \[ [Nₐ] = \lambda s \lambda x.x \text{ bears } Nₐ \text{ in } s = Pₐ(x)(s) \text{ if } Pₚ(x)(s), \text{ undefined otherwise.} \]

We further suggest that some nouns are *ambiguous* between a common noun use and a quasi-proper noun use, such as *haauzoeng* [C] or *headmastar* [B] ‘principal’:

(17)

a. \[ [\text{haauzoeng}_c/\text{headmastar}_c] = \lambda s \lambda x.\text{principal}(x) \] (common noun)
   b. \[ [\text{haauzoeng}_q/\text{headmastar}_q] = \lambda s \lambda x.\text{haauzoeng}/\text{headmastar}(x)(s) \text{ if } \text{principal}(x)(s), \text{ undefined o/w.} \]

The set of ambiguous nouns is language-specific and depends on naming convention.¹⁰

It can be diagnosed by vocative and title uses of human nouns. Those quasi-proper nouns can always be used in vocatives and titles, unlike common nouns:¹¹

---

⁸ Common nouns in classifier languages are usually treated as kind-denoting (Krifka 1995; Chierchia 1998; Yang 2001; Jiang 2020). For simplicity, we follow Trinh (2011) and assume a property denoting analysis, but we stress that our proposal is compatible with a kind-denoting treatment.

⁹ The name-bearing property may be formalized as a set of referential indices conventionally associated with the name by a set of assignment functions (see the discussion in Muñoz 2019).

¹⁰ One motivation for recruiting a common noun form as a name (i.e. a quasi-proper noun) could be to avoid addressing the referent’s proper name directly (see the discussion in §6).

¹¹ Note that the reverse does not apply: not all human nouns with the vocative and title usages have a quasi-proper noun counterpart, such as *tunghok* ‘schoolmate/student’ in Cantonese (vs. *hoksaang...
(18)  a.  \(\text{Wei, haauzoeng/*hoksaang, ngo soeng man je!}\)  (Vocatives)
   \(\text{principal/student 1SG want ask thing}\)
   Literally: ‘Hey, Principal/*Student! I have a question to ask.’  [C]

   b.  \(\text{Rashtropoti/*chhatro Washington}\)  (Titles)
   \(\text{president/*student Washington}\)
   ‘President Washington/*Student Washington’  [B]

Before proceeding to the DP level, we treat classifiers in bare CLs as performing
an atomic check AT for the counting function, and have the type \(\langle\langle s,\langle e,t\rangle\rangle,\langle s,\langle e,t\rangle\rangle\rangle\) in (19) (after Trinh 2011, against
the iota account in Cheng & Sybesma 1999).

(19)  a.  \(\text{CL} = \lambda P\lambda x.\mathbf{[}P(x)(s)\mathbf{]}\) if \(x \in AT(P(s))\), undefined otherwise.

   b.  \(AT(P(s)) = \lambda x.\mathbf{[}x \in P(s) \land \forall y[(y \in P(s) \land y \leq x) \rightarrow (y = x)]\mathbf{]}\)

On the DP level, we propose that there are two definite determiners: \(D_1\), and \(D_2\),
one for definite descriptions, another one for names. Definite bare CLs are derived
by \(D_1\). \(D_1\) is null in both languages, but it triggers syntactic movement (CL-to-D
movement in Cantonese, Simpson 2005; Wu & Bodomo 2009; NP-movement to
Spec DP in Bangla, Dayal 2012). As given in (20), \(D_1\) denotes an iota operator and
presupposes existence and uniqueness of the referent, like English the (Frege 1892;
Strawson 1950; Donnellan 1966, i.a.). We follow Schwarz (2009) in incorporating
a resource situation variable \(s_r\) to capture the situation-dependency of uniqueness.

(20)  The denotation of \(D_1\)-∅
      \[\mathcal{D}_1\] = \(\lambda P\lambda x.\mathbf{[}P(x)(s_r)\mathbf{]}\) if \(\exists! x[P(x)(s_r)]\), undefined otherwise.

Unique bare Ns (quasi-names) and proper names are derived by \(D_2\). \(D_2\) is also
null, but may be realized as aa- in Cantonese or -moshai in Bangla with a [+human]
NP. Let us give some prerequisites before spelling out \(D_2\). First, to formalize the
discourse, we adopt Kaplan (1977, 1989)’s context of utterance, as in (21).\(^{14}\)

(21)  \(\langle c_s, c_a, c_T, c_P, c_W\rangle\), where \(c_s, c_a, c_T, c_P, c_W\) are the speaker, addressee, time,
position, world of a context \(c\) respectively.

---

\(^{12}\) Or individuating kinds under a kind-denoting analysis (see Jiang 2020 and the references therein).

\(^{13}\) We remain open on whether English the is ambiguous between a weak article and a strong article that
carries an index, as proposed in Schwarz 2009 and Jenks 2018. What is important here is that \(D_1\)
patterns with the regarding uniqueness and anaphoric uses.

\(^{14}\) Kaplan originally has \(c_A\) as the agent (speaker) of \(c\), and does not have the addressee represented.
Are there “weak” definites in bare classifier languages?

Second, as discussed in §3, the functional relation is restricted to a set of relations that only apply to the individuals denoted by the nouns, for example, principal-of for ‘principal’.\(^\text{15}\) To capture it, we propose a \textit{functional characteristic relation} (FCR), on a par with Montague’s characteristic function (i.e., \textit{all} and \textit{only} the entities in the extension of a property \(P\) satisfy \(P\)), as defined in (22).

(22) A function \(f\) is a \textit{functional characteristic relation} (i.e. \(\text{FCR}(f)\)) iff:
\begin{enumerate}
\item \(f\) has the type \(\langle\langle s, e, t\rangle, \langle e, e\rangle\rangle\); and
\item for every non-empty predicate \(P\), there exists an entity \(z\) such that \(f(z)(P)\) is defined and it is in the extension of \(P\) (i.e., \(P(f(z)(P))\) holds); and
\item there does not exist an entity \(y\) such that \(y\) is not in the extension of \(P\) and \(f\) maps \(P\) and \(z\) onto \(y\) (i.e., \(\neg\exists y[\neg P(y) \land f(z)(P) = y]\) holds).
\end{enumerate}

\(f\) maps \(P\) to a relational function from entities to entities whose range is \(P\). In other words, \(f\) is similar to choice functions in giving an entity that satisfies \(P\) (Winter 1997), but the way of picking up involves a relation with some entities which does not hold for entities that do not satisfy \(P\). Effectively, only principal-of-\(z\) (or principal-of-\(z\)’s-school) but not mother-of-\(z\) is the relevant functional relation for a predicate ‘principal’ even if some principals happen to be someone’s mother.

We can now give the semantics of \(D_2\) in (23):

(23) The denotation of \(D_2\cdot\emptyset/\text{aa-}/\text{-moshai}\):

\[
[D_2]^{c,g}_r = \lambda s_r \lambda P_n. \begin{cases} 
\text{tx}[P_n(x)(s_r) \land g(i) = x \land } & \text{if } \exists! x [P_n(x)(s_r) \land g(i) = x] , \\
\exists f[\text{FCR}(f) \land f(c_a \oplus c_a)(P_n) = x] , & \text{undefined,} \\
\text{otherwise.} & \end{cases}
\]

We follow Muñoz (2019) and Agolli (2023) in assuming \(D_2\) to carry an index taken by a world-insensitive assignment function to map onto the referent, which gives the rigid designation of names. Moreover, \(D_2\), like \(D_1\), also denotes an \textit{iota} operator, but it additionally encodes an FCR function that maps a naming-bearing property \(P_n\) and the discourse participants (speaker \(c_a\) and addressee \(c_a\)) to the referent. The \(P_n\) taken by \(D_2\) restricts the combining nouns to be (quasi-)proper nouns (after Muñoz 2019). \(D_2\) presupposes that there is exactly one individual that (i) bears the relevant name \((P_n)\) in a resource situation \(s_r\), and (ii) is the value of \(j\) that maps from \(P_n\) and the sum of \(c_a\) and \(c_a\) in a context \(c\), which is the intended referent \((g(i) = x)\). For ease of reference, we call the presuppositions as \textit{Uniqueness} (the shared \textit{iota} part with \(D_1\)) and \textit{Relation} (the underlined part, not shared with \(D_1\)).\(^\text{16}\)

\(^{15}\) The set cannot be specified as inherent relations of relational nouns (see §3). Moreover, as discussed below, \(D_2\) takes (quasi-)proper nouns, which denote name-bearing properties and are not relational.

\(^{16}\) We are aware that the index in \(D_2\) can never be bound (\textit{cf.} Agolli 2023), unlike other definite expressions with an index (e.g. pronouns, demonstratives, and anaphoric definites). A potential
With the proposed entries of nouns and determiners, we can now give a semantic composition for bare CLs and bare Ns. First, definite bare CLs, such as `go-haauzoeng/headmastar-Ta` ‘the principal’ with the LF structure in (24a), denote a regular definite description in (24b). The resulting definite description, just like English *the NP*, can be used in both unique and anaphoric contexts (cf. footnote 13). Note that *s_r* can be supplied by the context via a topic situation, or by the situation quantified by other operators such as ‘every time’ to give a co-varying reading (i.e., each situation has one unique referent) (Schwarz 2009).

(24) a.  
\[
\begin{array}{c}
\text{[DP D}_1-\emptyset [\text{CLP CL-go/Ta [NP N-haauzoeng/headmastar}}_c ]] \\
\text{[DP]}^{c, g}_x = \text{tx.}\text{[principal}(x)(s_r)} \\
\text{if } \exists ! x[\text{principal}(x)(s_r)] \land x \in \text{AT}([\text{principal}(s_r)], \text{undefined otherwise}) \\
\text{[D}_1-\emptyset ]^{c, g}_x = \lambda Px[P(x)(s_r)] & \text{if } \exists ! x[P(x)(s_r)] \\
\text{[CLP]}^{c, g} = \lambda s\lambda x.\text{principal}(x)(s) & \text{if } x \in \text{AT}([\text{principal}(s)], \text{undefined otherwise}) \\
\text{[NP-haauzoeng/headmastar}_c ]^{c, g} = \lambda s\lambda x.\text{principal}(x)(s)
\end{array}
\]

b.  
\[
\begin{array}{c}
\text{[DP]}^{c, g}= \text{tx.}\text{[principal}(x)(s_r)} \\
\text{if } \exists ! x[\text{principal}(x)(s_r)] \land x \in \text{AT}([\text{principal}(s_r)], \text{undefined otherwise}) \\
\text{[D}_1-\emptyset ]^{c, g}_x = \lambda Px[P(x)(s_r)] & \text{if } \exists ! x[P(x)(s_r)] \\
\text{[CLP]}^{c, g} = \lambda s\lambda x.\text{principal}(x)(s) & \text{if } x \in \text{AT}([\text{principal}(s)], \text{undefined otherwise}) \\
\text{[NP-haauzoeng/headmastar}_c ]^{c, g} = \lambda s\lambda x.\text{principal}(x)(s)
\end{array}
\]

Now we turn to bare Ns like *haauzoeng/headmastar* ‘the principal/Principal’. With the structure in (25a), it denotes a referential quasi-name as in (25b), which can be used in a subset of uniqueness contexts where there is a unique individual in a given situation that bears the name “Principal” with which the discourse participants hold a relation, understood as principal-of. Due to the rigid designation, the referent is always anchored to the root context even if *s_r* is quantified by other operators.\(^{17,18}\)

---

17 \footnote{Even if there is a relation between the discourse participants and every unique referent in the quantified situations, quasi-names still have no co-varying readings in ‘Every time we went back to our primary and secondary schools, Principal came to welcome us’ (assuming both schools have different principals), like proper names. We thank Florian Schwarz for bringing up this possibility.}

18 \footnote{Quasi-names can be used in the presence of other individuals bearing such names as long as the speaker/addressee do not have a relation with them, such as *Mom* in English as in *Mom is talking to other moms*. We thank Veneeta Dayal for pointing this out. Note also that there could well be two individuals with which speaker/addressee have a relation, such as *Grandma*. Quasi-name *Grandma* can only be used if there is one grandma in the given situation, and is infelicitous when both grandmas are present (where one needs to resort to other means like ‘maternal grandma’ or proper names). We thank Florian Schwarz for bringing up this case.}
Are there "weak" definites in bare classifier languages?

(25) a. \[ \text{DP } D_2 \rightarrow \emptyset / aas / moshaiz [\text{NP } haauzoeng_q / headmastar_q (=h_q)] \]

\[ [\text{DP}]^{\text{c-e}} \]
\[ = \lambda x [h_q(x)(s_f) \land g(5) = x \land \exists f [FCR(f) \land f(c_a + c_a)(h_q) = x]] \]
\[ \text{if } \exists! x [h_q(x)(s_f) \land g(5) = x \land \exists f [FCR(f) \land f(c_a + c_a)(h_q) = x] \land \text{principal}(x)(s_f), \text{u.o.} \]

b. \[ [D_2 \rightarrow \emptyset]^{\text{c-e}} (s_f) \quad [\text{NP-}haauzoeng_q / headmastar_q]^{\text{c-e}} \]
\[ = \lambda s \lambda P_a. \lambda x [P_a(x)(s_f) \land g(5) = x \land \exists f [FCR(f) \land f(c_a + c_a)(P_a) = x]] \quad \text{if } \exists! x [P_a(x)(s_f) \land g(5) = x \land \exists f [FCR(f) \land f(c_a + c_a)(P_a) = x]] \]
\[ = \lambda s \lambda x. h_q(x)(s) \quad \text{if } \text{principal}(x)(s) \]

Lastly, we note that D_2 is not stipulated for unique bare Ns, but it is a proprial article that may combine with proper nouns. Adopting predicativism of names (e.g., Elbourne 2005; Agolli 2023), the entry of D_2 is independently needed for the interpretation of referential proper names. This explains why unique bare Ns exhibit a striking parallelism with proper names.\(^{19}\)

6 Competing referring expressions

In this section, we turn to competition among referring expressions. This is common in natural languages and has been accounted for by economy principles (Heim 1991; Jenks 2018; Ahn 2019, 2023; Dayal & Jiang 2022; Owusu 2022, i.a.).\(^{20}\) We have seen some cases where bare Ns are chosen over bare CLs because of the existence of a functional relation (cf. the Relation presupposition in D_2). In the following, we further elaborate on the exact mechanisms that drive the competition between quasi-names (bare Ns) and definite descriptions (bare CLs), and an understudied case of quasi-names competing with proper names. A preview is given below.

(26) Case #1. ✔ Bare CLs vs. ✗ Bare Ns (only bare CLs)
Case #2. ✗ Bare CLs vs. ✔ Bare Ns (only bare Ns)
Case #3. ✔ Bare CLs vs. ✔ Bare Ns (both bare CLs and bare Ns)
Case #4. ✗ Bare CLs vs. ✗ Bare Ns (neither bare CLs nor bare Ns)

In the first case, only bare CLs are allowed. It involves a context where only the Uniqueness presupposition is met, but the Relation presupposition is not, which is the scenario in the anaphoric cases (e.g., referring back to a referent in an antecedent

\(^{19}\) This amounts to saying that some characteristic relations also hold for the referent denoted by proper names (say Smith), which we take to be a call-x-by-x’s-name relation, i.e., the discourse participants call the referent by Smith. This captures that the use of proper names requires some knowledge about the referent so as to identify them by their name (Prince 1992; Abbott 2002).

\(^{20}\) These economy principles may be driven by pragmatic forces in different directions. For example, Heim’s (1991) Maximize Presupposition! and Jenks’s (2018) Index! appear to be driven by a need to be more informative (cf. Horn’s (1984) Q Principle), whereas Ahn’s (2019) Don’t Overdetermine! (or Minimize Restrictors! in Ahn 2023) discourages over-information (cf. Horn’s R Principle).
clause in (1)) and uniqueness cases where no relation is held between the discourse participants and the referent (e.g. ‘the principal’ in another school in (3) and ‘the (alien) moon’ in (5)). In these cases, using $D_2$, which presupposes both Uniqueness and Relation, would lead to presupposition failure, and only $D_1$ can be used.

In the second case, bare Ns are used and bare CLs are blocked. It involves a scenario where both Uniqueness and Relation presuppositions are satisfied (e.g. ‘the principal’ in the same school in (2) and ‘the (earth) moon’ in (4)). The choice of the bare N is in accordance with the principle of Maximize Presupposition! (Heim 1991; Sauerland 2003, 2008), which states that if the two forms contribute the same assertive component relative to the conversational context, the form with a stronger presuppositional strength blocks the form with a weaker one. While bare Ns and bare CLs contribute to the same assertive component (i.e., denoting the referent), $D_2$ in bare Ns carries a stronger presupposition (both Uniqueness and Relation) than $D_1$ in bare CLs (only Uniqueness). Hence, the stronger form $D_2$ is chosen over $D_1$.

In the third case, both bare CLs and bare Ns are felicitous, as exemplified in (27):

(27) **Context:** A music band constituted of school students is visiting a different school for a music competition. They have been waiting for a long time as the show can’t start before the principal arrives. A member of the band says:

a. \{Go-haauzoeng/ haauzoeng\} zung mei dou.
   \textsuperscript{cl-principal principal} still not.yet arrive
   ‘The principal hasn’t arrived yet.’ \[C\]

b. \{headmaster-Ti/ headmaster\} ekhono ashe ni.
   \textsuperscript{principal-cl principal} still come NEG
   ‘The principal hasn’t arrived yet.’ \[B\]

The discourse participants in (27) are not part of the school and thus the referent is not their principal. This is similar to the first case where only Uniqueness is met and only bare CLs can be used. Nevertheless, the bare N is also felicitous in (27). We suggest that this case involves perspective shift followed by presupposition accommodation (von Fintel 2008). In (27), the topic situation is the school, which enables the speaker to pretend to hold a functional relation with the principal, and such a Relation presupposition is accommodated by the addressee. Therefore, $D_2$ can be used in addition to $D_1$.\textsuperscript{22}

In the last case, neither bare CLs nor bare Ns are allowed. Instead, a proper name or a possessive construction is preferred, as illustrated in (28):

\textsuperscript{21} See Schlenker (2012: 392-393) for the definitions of assertive component and presuppositional strength.

\textsuperscript{22} Accommodation also happens in some cases where the relation is unclear. In an example like (i), both the bare N and bare CL are felicitous even though the discourse participants are not related to the driver in a typical sense (Andrew Simpson p.c.).
Are there “weak” definites in bare classifier languages?

(28)  Context: You and your family are sharing various stories about your pet dog Bobby. Some of them portray that the dog is silly. You, on the other hand, want to defend your dog. You say:

a. {#Zek-gau/ #gau/ ngodei zek-gou/ Bobby} hou gwaai gaa3.  
   CL-dog  dog  IPL  CL-dog  Bobby  very well-behaved SFP  
   ‘(Whatever you say) the dog/ Our dog/ Bobby is well behaved.’  
   [C] 

b. {#kukur-Ta/ #kukur/ ama-der kukur/ Bobby} kintu khub-e bhodro.  
   dog-CL  dog  IPL-GEN dog  Bobby  but very-INT well behaved  
   ‘(Whatever you say) the dog/ Our dog/ Bobby is very well-behaved.’  
   [B] 

In (28)’s context, both Uniqueness and Relation presuppositions are met (i.e., the dog is “our” dog), similar to the second case where only bare Ns can be used. The infelicity of bare CLs is expected from Maximize Presupposition! which chooses D2 over D1. Yet, bare Ns are also disallowed. We suggest that it is due to a ‘name competition’ between quasi-names and proper names at the NP level: the existence of the proper noun ‘Bobby’ blocks the quasi-proper noun use of ‘dog’. Unlike the principal cases where there is a conventional force to prefer a quasi-name over a proper name like avoiding directly calling the principal’s names out of courtesy, no such forces are present in the ‘dog’ case. There is thus no recruitment of the noun form ‘dog’ as a name, and the quasi-proper noun use is blocked. Since both quasi-names and proper names are formed by D2, the competition is only at the NP level. Another choice is to use a common noun instead of a proper noun, which is the case of possessive constructions with the structure [PossP our [NP Common Noun ]]. The common noun does not compete with the proper noun: it denotes a descriptive property and has a different assertive content. The first-person plural possessive also captures the functional relation. From this case, we see a competition that occurs at both DP levels (D1 vs. D2) and NP levels (quasi-proper noun vs. proper noun).

7 Concluding remarks

To conclude this paper, we have argued that definite bare Ns in Cantonese and Bangla are quasi-names but not unique definites. Definite bare CLs are not anaphoric.
definites either - they are standard definite descriptions, covering both uniqueness and anaphoric uses. We have offered a compositional analysis that builds on different types of nouns (descriptive property vs. name-bearing property) at the NP level and determiners (D₁ and D₂) at the DP level. Notably, we have proposed that D₂ is a proprial article that encodes a functional relation between the discourse participants and the referent to form both quasi-names and proper names. We also discussed the competition between bare Ns vs. bare CLs, and proper names vs. quasi-names.

The findings have implications for both the typology of definiteness and competition among referring expressions. **First**, we need a rigorous re-examination of the typology of definites that integrates (quasi-)names, particularly for bare Ns that have been analyzed as unique definites. Indeed, we are not the first to suggest that bare Ns are name-like expressions, and similar ideas have been alluded to by Cheng & Sybesma (1999) and Jenks (2018) for Cantonese, as well as Akan (Bombi, Grubic, Renans & Duah 2019, but see Owusu 2022), Mandarin (Cheng & Sybesma 1999), and Hindi (cf. Bhatt & Davis 2023).²³ A new typology with quasi-names is given in Table 3. **Second**, this new typology also informs us about the competition: names, just like demonstratives and pronouns (Jenks 2018; Ahn 2019, i.a.), also compete with definite descriptions. The competition may occur at two levels: the DP level (e.g. D₁ vs. D₂ modulo Maximize Presupposition!), and the NP level (e.g. quasi-proper noun vs. proper noun), the latter of which resonates Ahn (2019, 2023)’s economy principles in regulating restrictors (= NP). This sheds further light on how these economy principles, which may be driven by pragmatic forces in different directions, operate at different levels to determine the form of referring expressions.

<table>
<thead>
<tr>
<th>Language</th>
<th>Type</th>
<th>Definite description</th>
<th>Quasi-names</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cantonese</td>
<td>CL-lang.</td>
<td>bare CL</td>
<td>bare N/aa-</td>
</tr>
<tr>
<td>Bangla</td>
<td>CL-lang.</td>
<td>bare CL</td>
<td>bare N/-moshai</td>
</tr>
<tr>
<td>Mandarin</td>
<td>CL-lang.</td>
<td>bare N</td>
<td>bare N</td>
</tr>
<tr>
<td>Hindi</td>
<td>Non-CL-lang.</td>
<td>bare N</td>
<td>bare N/-ji</td>
</tr>
<tr>
<td>Akan</td>
<td>Non-CL-lang.</td>
<td>determiner no</td>
<td>bare N</td>
</tr>
<tr>
<td>German</td>
<td>Non-CL-lang.</td>
<td>weak art. strong art.</td>
<td>?</td>
</tr>
<tr>
<td>Fering</td>
<td>Non-CL-lang.</td>
<td>weak art. strong art.</td>
<td>?</td>
</tr>
</tbody>
</table>

**Table 3** The typology of definiteness with quasi-names (final)

²³ Bhatt & Davis (2023) do not explicitly mention this, but they show that the honorific name-marking device -ji can attach to bare Ns.
Are there “weak” definites in bare classifier languages?

References


Are there “weak” definites in bare classifier languages?


Are there “weak” definites in bare classifier languages?

Ka-Fai Yip  
Dow Hall, 370 Temple St  
Yale University  
New Haven, CT 06511  
kafai.yip@yale.edu

Ushasi Banerjee  
Dow Hall, 370 Temple St  
Yale University  
New Haven, CT 06511  
ushasi.banerjee@yale.edu

Margaret Chui Yi Lee  
Room 368 Oak Hall  
University of Connecticut  
Storrs, CT 06269-1145  
chui_yi.lee@uconn.edu