(In)definiteness without Articles: Diagnosis, Analysis, Implications

Veneeta Dayal – Rutgers University
In Ghanshyam Sharma and Rajesh Bhatt (eds), *Trends in Hindi Linguistics*, Mouton de Gruyter, Berlin: 2018

Abstract. This paper addresses the view that bare nominals in languages without articles can be definites as well as indefinites. In particular, it challenges the status of bare nominals as indefinites, providing contexts in which English uses the indefinite article but an article-less language like Hindi must resort to a numeral construction. The empirical generalization based on Hindi is that bare nominals are ambiguous between definites and kind terms and that bare plurals, but not bare singulars, can have kind derived indefinite readings. The indefinite readings available to bare singulars must be traced to external factors. The paper then provides an explanation for these facts within a Neo-Carlsonian approach to bare nominals, addressing the issue of null determiners vs covert type shifts. The last section of the paper extends the insights gained from the case study of Hindi bare nominals to other languages, highlighting empirical and theoretical issues of relevance in determining whether bare nominals in a given language are definites or indefinites. It also comments on the well-established fact that adult learners of a language like English have trouble with the determiner system if their L1 does not have articles from the vantage point of the findings about Hindi bare nominals.

1. The Definite-Indefinite Dichotomy

In this section we introduce two perspectives on the (in)definiteness of bare nominals in article-less languages, one which holds them to be ambiguous between definite and indefinite and one which takes them to be indefinites with a wider distribution. I argue against both perspectives by looking closely at bare nominals in Hindi. On the basis of explicit diagnostics it is established that bare nominals can be definites as well as kind terms. Bare plurals, but not bare singulars, can also have narrow scope indefinite readings that can be ascribed to their status as kind terms. Other possible sources for indefinite readings, available to both bare singulars and bare plurals, are derivative on properties of the constructions they appear in.

1.1. The Ambiguity of Bare Nominals

That bare nominals in languages without articles can function as definites as well as indefinites is a reasonable assumption when viewed from the perspective of languages with articles. The following comment from Löbner (1985:320) is representative of this view: “as for languages which do not have a definite article, it is plausible to assume that they just do not explicitly

1 I would like to thank the audiences at the Workshop on Semantic Universals (University of Pennsylvania), TripleA 3 (University of Tuebingen) and University College London for helpful comments. All errors and omissions remain my responsibility. Parts of this paper were included in the *Proceedings of Triple A*. 
express the way nouns are to be interpreted.” I take this to imply that bare nominals in such languages are ambiguous between definite and indefinite.

A more nuanced view is presented in Heim (2011): “in languages without definiteness marking, the relevant “ambiguous” DPs may simply be indefinites. They are semantically equivalent to English indefinites. But they have a wider range of felicitous uses than English indefinites, precisely because they do not compete with definites and therefore do not get strengthened to carry the implicatures that would show up if they were uniformly translated as indefinites into English”. To put this in context, consider the following meanings for the definite and indefinite articles. (1a) encodes the uniqueness of singular definites, while (1a’) gives an extension that is applicable to both singular and plural definites.2

\[
\begin{align*}
(1a) & \quad [\text{the}] = \lambda P : \exists x \forall y [P(y) \leftrightarrow y = x]. \iota x[P(x)] \\
(a') & \quad [\text{the}] = \lambda P : \exists x \forall y [(\text{MAX}(P))(y) \leftrightarrow y = x]. \iota x[(\text{MAX}(P))(x)] \\
& \quad \text{Where } \text{MAX}(P) = \lambda x [P(x) \& \neg \exists y [P(y) \& x < y]] \\
(b) & \quad [\text{a/an}] = \lambda P. \lambda Q. \exists x [P(x) \& Q(x)] \\
(c) & \quad a/an \quad \text{the}
\end{align*}
\]

The definite and indefinite articles differ in two respects. The former presupposes uniqueness -- it can be used felicitously only in contexts where there is exactly one individual that satisfies the common noun predicate -- and it picks out that unique individual as its referent. The latter imposes no requirements on how many individuals must satisfy the common noun predicate. It denotes a generalized quantifier, namely the set of properties such that at least one individual who has the common noun property also has that property. Focusing on the first difference, we see that it sets up an entailment relationship between the definite and the indefinite. Every context in which a statement with the N is felicitous and true is a context in which the corresponding statement with a/an N is also felicitous and true, but not the other way around.

Consider from this perspective a context in which it is known that there is exactly one cat who may or may not be sitting on the mat.

\[
\begin{align*}
(2a) & \quad \text{The cat/#A cat is on the mat.} \\
(b) & \quad \text{The sun/#A sun is shining.}
\end{align*}
\]

Both sentences in (2a) are defined in this context – whether they are true or false would depend on what that cat is doing. The version with the indefinite has an implicature that

---

2 For present purposes we work with the simpler version in (1a). The version in (1a’) includes plural individuals and imposes uniqueness on maximal individuals, those that are not individual parts of bigger individuals. In the case of singular common nouns P = MAX(P), but in the case of plural individuals MAX(P) is a set with only one individual, the one that is the sum of all the singular individuals in P. This facilitates the shift from the perceived uniqueness of singular terms (the man is tall) to the perceived universality of plural terms (the men are tall).
there should be more than one cat in the context. This is due to competition with the
definite since Gricean principles of conversation require a speaker to use the strongest
statement compatible with the known facts. The perceived infelicity of the indefinite in
(2a) is thus accounted for. This is illustrated even more dramatically in (2b) where world
knowledge entails the existence of a single sun.

Heim’s statement, that the so-called ambiguity of bare NPs in languages without
articles arises simply from lack of competition, can be schematically represented in (3a).
The more standard view is represented in (3b):

(3a) ![Diagram of bare NPs in article-less languages]

(3b) ![Diagram of bare NPs in article-less languages]

These two views about the interpretation of bare nominals in articleless languages are
echoed in the literature on Hindi, which we will take as our primary example of this type of
language. They have been taken to be ambiguous between definite and indefinite (Mahajan
1990, Mohanan 1994, for example), and as indefinites whose definite interpretation is
pragmatically derived (Kidwai 2000 and Thakur 2015). The next subsections will show
that Hindi bare nominals are indeed definites but will provide evidence to challenge the
claim that they are indefinites. In effect, then, we will be arguing also against the view that
Hindi bare nominals are ambiguous between specific and definite meanings (Verma 1971,
Masica 1991), and asserting the claim in Kachru (1980) and Dayal (2004) and previous
work that they are ambiguous between definites and generics only.3

1.2. The Definiteness of Bare Nominals

There are two clear diagnostics that together determine whether a bare noun phrase is definite:
anaphoricity and homogeneity. Let us consider the anaphoricity test first:

(4a) A boy and a girl came into the room. The girl sat down.
(b) ek laRkaa aur ek laRkii kamre meN aaye. laRkii baith-gayii.
    one boy       and one girl      room   in     came. girl     sat-down
    ‘A boy and a girl came into the room. The girl sat down.’

3 Thanks to Deepak Alok for these references.
As we see in (4a), a definite is required in the second sentence to refer back to a discourse referent introduced in the first sentence. An indefinite would not be felicitous. When we translate the discourse into Hindi in (4b), we see that the bare NP laRkii ‘girl’ is quite felicitous.\(^4\) Thus the bare NP passes the first test for definiteness.

The homogeneity test, due to Löbner (1985), differentiates definites from demonstratives as well as indefinites, and aligns it with proper names. This is in keeping with the semantics of the definite determiner given above:

\[
\begin{align*}
(5a) & \quad \text{A dog is sleeping and a dog is barking.} \\
(b) & \quad \text{That dog is sleeping and that dog is barking.} \\
(c) & \quad \#\text{The dog is sleeping and the dog is barking.} \\
(d) & \quad \#\text{Fido is sleeping and Fido is barking.}
\end{align*}
\]

Given the analysis of definites as denoting a single individual, it follows that one cannot predicate incompatible properties of a definite. That is, they behave in this respect like proper names. The test also applies to plural definites (#the dogs are sleeping and the dogs are barking), though the test has to be administered with some care.\(^5\) Turning to Hindi, we see that a bare NP in this context is unacceptable:

\[
\begin{align*}
(6) & \quad \text{kutta so-rahaa-} & \text{aur } \text{kutta bhaunk-rahaa-} & \text{hai.} \\
\text{Dog } & \text{sleep-PRG-PR and dog } & \text{bark-PRG-PR} \\
\text{’The dog is sleeping and the dog is barking.’}
\end{align*}
\]

Based on these two tests, then, we can safely classify the Hindi bare NP as a definite.

\subsection{1.3. The Non-indefiniteness of Bare Nominals}

The fact that bare NPs in languages without articles are definites is not controversial. But we will now look at diagnostics that call their status as indefinites into question. In fact, the example in (6) is already suggestive. If the bare NP kutta ‘dog’ could be interpreted as an indefinite in (6), the sentence would have been acceptable under the reading that the English (5a) has. The next three diagnostics drive home this point: partitive specificity, referential specificity, scopal inertness, all show the Hindi bare NP diverging from the English indefinite.\(^6\)

\(^4\) Using two distinct noun phrases in the first sentence makes the use of a full DP in the second sentence more natural by making a null or overt pronoun either unavailable or less preferred.

\(^5\) Briefly, plural definites do not require that every member of the set be actively involved. For example, the statement the reporters asked questions may be considered true even in contexts where only some of them did. However, it is still the case that the predication applies to the group as a whole so that the reporters asked questions but the reporters kept quiet or the reporters asked questions but some of them kept quiet would be deemed contradictory (see Dayal 2013 and references cited there).

\(^6\) See von Heusinger (2011), among others, for a survey of issues related to indefiniteness and specificity.
Consider the test of partitive specificity, discussed in Enç (1991), where the boy and girl referred to in the second sentence are understood as belonging to the set mentioned in the first:

(7a) There were several kids in the room. A boy and a girl were playing cards
(b) kamre meN kaii bacce the. #LaRkaa aur laRkii taash khel-rahe-the
   room in several kids were boy and girl cards play-PROG-PAST
   ‘There were several children in the room. #The boy and the girl were playing cards.’

No matter how we set up the context, if there were several kids in the room, there must be a plurality of girls or boys or both. If the bare NPs were indefinites, there would be no problem in getting a partitive specific reading, in analogy with (7a). But what we see in (7b) is the infelicity that is predicted of NPs that encode uniqueness. Given that both NPs have singular morphology, the context has to be one with a unique boy and a unique girl. In the given context, the numeral ek ‘one’ would have to be added to the NPs to convey the intended meaning.

Next consider referential specificity, brought into the public discussion by Fodor and Sag (1982). The important point here is that the speaker has a particular individual in mind, one that the hearer does not necessarily know about. Had the intended referent been identifiable to both the speaker and the hearer, the plain definite my relative instead of a relative of mine would have been used in (8a).

(8a) If a relative of mine dies, I will inherit a fortune.
(b) agar mere rishtedaar ki maut ho jaaye, to mujhe bahut paisaa milegaa
     if my relative of death happen then I-DAT lot money get-FUT
     ‘If my relative dies, I will get a lot of money.’

The bare NP in (8b) denotes a unique referent identifiable to both speaker and hearer. The addition of a numeral mere ek rishtedaar ‘my one relative’ is needed to make it parallel to (8a). It should be noted that this diagnostic is known to be a delicate one. The difference between a definite and specific indefinite with respect to referential specificity is slight, riding crucially on a gap between a speaker’s presupposition of uniqueness vs. common ground knowledge.\(^7\)

The distinction between bare NPs and indefinites becomes clearer when we look at cases of intermediate scope, by now a well-established fact in the literature on specific indefinites (Farkas 1981, among others). The intermediate scope reading in (9a) allows the choice of topic to vary with students. Thus, (9a) can be continued with Mary has read every article on fake news and Sue has read every article on the Affordable Care Act. This reading is distinct from the wide scope reading we saw in (8) as well as a narrow scope reading. A narrow scope reading for the

---

\(^7\) One may argue, quite rightly, that the definiteness in this example is due to the possessive construction. The absence of the relevant reading can be ascertained by translating examples like if a tall dark-haired student comes, tell her to wait which do not involve possessives. See von Heusinger (2011) for the types of modification that facilitate this reading.
indefinite would require every student to have read every paper on any current topic – that is every student would have read every paper on every current topic.\(^8\)

\[(9a)\quad \text{Every student has read every article on a current topic.}\]
\[(b)\quad \text{Every student-ERG current topic on every article read}\]
\[(c)\quad \text{Every student-ERG one/some-one current topic on every article read}\]

We see from (9b)-(9c) that Hindi needs an overt marker of indefiniteness to get the relevant reading, one that allows a follow-up with different students paired with different topics. Other examples can be constructed along the same lines to show that the surviving reading of a bare nominal is that of a definite:

\[(10a)\quad \text{Every teacher wrote two articles which he gave to a student of his.}\]
\[(b)\quad \text{Every teacher-ERG two articles wrote that he self’s student-DAT gave}\]

While (10b) allows students to vary with teachers, a reading that is forced by the anaphoric possessive in the relative clause, the presupposition of uniqueness projects. Each teacher is assumed to have a single student who is contextually relevant. What we do not get is the specific indefinite reading. In order to get that reading, we would need to add the overt marker of indefiniteness \text{apne ek vidyaarthii-ko} ‘self’s one student-DAT’. With this addition, each professor could have multiple students but only one of whom would be referred to as the recipient of the articles.\(^9\)

It turns out that the inability to take intermediate scope is simply a reflex of an apparently indefinite bare nominals’ inability to take scope over any clause-mate operator. In this respect Hindi bare nominals behave like English bare nominals:

\[(11a)\quad \text{John didn’t read a book/books.}\]
\[(b)\quad \text{Rabbits/A rabbit kept dying.}\]

\(^8\) It is unclear to me why some form of modification facilitates the relevant reading in (9a) with the indefinite article \text{a}, but is not needed with \text{some}. In the interest of space I am setting aside many important differences between different indefinite articles (see Farkas and Brasoveanu \textit{forthcoming}, among others, for relevant discussion). Hindi (9b) is not fully acceptable without modification. To get the definite reading without modification, a demonstrative \text{us vishai} ‘that topic’ seems to be required. At this point, I must content myself with simply noting this fact.

\(^9\) Again, we may want to steer clear of the possessive construction to ensure that it is not the locus of definiteness (see footnote 6). However, in this case, having the bare \text{vidyaarthii ‘student’ shifts the interpretation to one where there is exactly one student assumed to be salient in the discourse. The reason for the missing narrow scope reading will be taken up in section 1.4.
Dogs have evolved from wolves.

As originally pointed out by Carlson (1977), a sentence like (11a) with a singular indefinite has two readings, one where there is a particular book John didn’t read while he may have read others ($\exists > \neg$); and another in which he read no books at all ($\neg > \exists$). The bare plural has only the latter reading. This may suggest that bare plurals are a sub-type of indefinites but the example in (11b) shows that this is not the case. The indefinite has the implausible reading where the same rabbit died several times ($\exists > \text{Duration}$) while the bare plural has a plausible differentiated reading that allows for different rabbits to die at different times ($\text{Duration} > \exists$). Carlson’s account of bare plurals is that they are not indefinites at all but rather kind terms and that they can be arguments of kind level predicates, as in (11c). When they occur as arguments of individual and stage level predicates, semantic operations make their instantiations available for semantic composition. We will come back to details of the theory in section 2. For now it is important to note that there is a correlation between scopal inertness and kind terms, a correlation that is also present in Hindi:

(12a) sher vilupt ho jaa saktaa hai/sher vilupt ho jaa sake haiN
tiger extinct become can tigers extinct become can
‘The tiger can become extinct.’/ ‘Tigers can become extinct.’

(b) tiin din tak
three days till
‘For three days’
khargosh marte rahe/#khargosh martaa raahaa/#ek khargosh martaa raahaa
rabbits die-PROG/#rabbit die-PROG/ one rabbit die-PROG
‘Rabbits kept dying.’/ ‘The rabbit kept dying’/ ‘A rabbit kept dying.’

In (12a) we have a kind-level statement with a bare singular and a bare plural. In (12b) we see the differentiated scope behavior of bare plurals, where different rabbits can be understood to die during the interval referred to. We note that the same is not true of the bare singular or the indefinite.¹⁰ We will return to this difference between bare plurals and bare singulars, both of which are bona fide kind terms, in the next subsection.

What we have established here is that Hindi bare nominals are not like English indefinites. In particular, they cannot be classified as specific or partitive indefinites. We have also noted that they are scopally inert and related it to their being kind terms. Interestingly, this scopal inertness results in readings, at least in the case of bare plurals, that are unavailable to regular indefinites. We will now turn to a consideration of the evidence that has formed the basis of the commonly held view that bare nominals in languages like Hindi are indefinites.

¹⁰ Note that a bare singular does not lend itself to a narrow scope reading in conditionals. The Hindi translation of if a student filed a complaint, she will be reprimanded would need koi vidyaarthii ‘some student’. The bare singular would suggest the existence of a single contextually salient student, as mentioned in footnote 8.
1.4. The Pseudo-indefiniteness of Bare Nominals

One of the strongest motivations for taking bare nominals to be indefinites comes from examples like (13a). A background question like *what were you doing yesterday afternoon?* removes the likelihood of there being a salient book in the common ground. And answers like (13a) in such a context are most naturally translated with an indefinite determiner. However, this does not necessarily mean that the indefiniteness is inherent to the bare nominal, as (13b)-(13c) show:

(13a)  *maiN kitaab paRh-rahii-thii*
       I book read-PROG-PAST
       ‘I was reading a book.’

(13b)  *maiN bacce-ko khaanaa khilaa-rahii-thii*
       I child-DAT food feed-PROG-PAST
       ‘I was feeding the child.’

(13c)  *maiN baccoN-ko khaanaa khilaa-rahii-thii*
       I children-DAT food feed-PROG-PAST
       ‘I was feeding children/the children.’

If the person answering the question with (13b) or (13c) works in a daycare or any situation with multiple children in it, there is no reason why the bare singular indirect object should not have an indefinite reading. The fact that (13c) can have a non-specific indefinite reading establishes that there is no pragmatic obstacle involved. What we can conclude from this paradigm is that the indefinite readings we see in (13) come from two distinct sources. One is the indefiniteness associated with plural kind terms (13c), the other is the indefiniteness associated with incorporation and/or complex predicate formation, a process that targets direct objects (13a). If indefiniteness were inherent to the bare nominal, changing from direct to indirect object position in (13b) would not make the difference it does.

A second case where indefiniteness is perceived is the existential construction. The position of arguments in relation to locative phrases has been noted to correlate with (in)definiteness. While there is clearly validity to this generalization, it is worth noting that it does not readily extend to all bare nominals, especially singular bare nominals:

(14a)  *cuuhaa kamre meN hai*
       mouse room in is
       ‘The mouse is in the room.’

(14b)  *kamre meN cuuhaa hai*
       room in mouse is
       ‘There’s a mouse in the room.’

(14c)  *kamre meN ??(ek) buuDhii aurat hai*
       room in one old woman is
       ‘There’s an old woman in the room.’

Why should the switch from *cuuhaa* ‘mouse’ to *buuDhii aurat* ‘old woman’ make a difference to indefiniteness? The bare plural counterpart of (14c) would be quite acceptable. This suggests
again that bare singulars in existential contexts are dependent on factors independent of quantificational properties of the noun phrase, factors that are not very well understood currently.

Let me end this section with a canonical property associated with indefinites, namely the ability to introduce discourse referents. We can test this by setting up a presentational context, ruling out the possibility of reference to familiar entities. Neither of the following seem acceptable without the addition of the unstressed numeral `ek’ `one’, specially if the story is expected to elaborate on the relevant individuals:

(15a) bahut saal pahle is ghar meN ?(ek) bhuuDhii aurat rahtii-thii
    Many year back this house in one old woman live-HAB-PAST
    ‘Many years ago, an old woman used to live in this house.’

(15b) bahut saal pahle is ghar meN ?(ek) vidyaarthii-kii maut ho-gayii-thii
    Many year back this house in one student-GEN death happen-PAST
    ‘Many years ago, a student had died in this house.’

Though there is some speaker variation regarding the ability of bare singulars to introduce novel discourse referents, I believe it can be safely claimed that for most speakers there are at least some contexts in which bare singulars cannot introduce new entities into the discourse. And for no speaker is it the case that bare NPs are preferred over the corresponding indefinite. Substituting bare plurals in (15a)-(15b) leads to marked improvement, making them perhaps even fully acceptable.

1.5. A Snapshot of the Empirical Landscape

In the previous subsections I have collected together the evidence in the literature establishing that bare nominals in an article-less language like Hindi are definites but not specific or non-specific indefinites and that their apparent indefiniteness is construction specific and/or restricted to bare plurals. I summarize the findings for indefiniteness in the following table:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Partitive Specificity</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Referential Specificity</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Wide/Intermediate Scope</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Narrow Scope Universal/ Negation</td>
<td>✓</td>
<td>✓/X</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Differentiated Scope under Adv</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Presentational</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
There are a couple of points worth noting in the above. We have not discussed the inability of the numeral NPs *ek N* in Hindi to take scope under negation. This will be taken up in section 3. I have also idealized the data for bare singulars in presentational contexts. As noted in section 1.4 this may well be contextually sensitive and subject to speaker variation. The point to note though is that bare nominals are not specific indefinites. And it is not at all clear that their non-specific indefinite behavior can be delinked from their life as kind terms. The fact that singular and plural terms behave differently with respect to differentiated scope and presentational contexts is significant in this connection. I therefore take the relevant generalization to be that bare nominals in Hindi are ambiguous between definites and kind terms. In section 2 I provide a theoretical account for this generalization. Those not particularly interested in the theory or already familiar with it should feel free to skip to section 3 where I discuss how the empirical landscape outlined above could be used to study the (in)definiteness of bare nominals in other languages as well as the acquisition of the determiner system by adult learners whose first language does not include determiners.

### 2. A Neo-Carlsonian Account of Bare Nominals

This section presents an account of the facts discussed in section 1. It recognizes the universality of operations that yield kind, definite and indefinite readings, allowing for these operations to have lexical exponents or be available as covert type shifts. The theory also includes constraints on type shifts, in relation to overt determiners and in relation to each other. Finally, a formal distinction between singular and plural kind terms is drawn with respect to individuals that at the intuitive level are instantiations of the kind. Together they account for the core generalization that bare nominals in article-less languages are bona fide definites but not bona fide indefinites. The section ends by addressing the implications of treating bare nominals as DPs with a null determiner versus NPs that undergo covert type-shift in argument positions.

#### 2.1. Bare Plurals as Kinds and Definites

The basic premise of the neo-Carlsonian approach is that bare plurals in argument position refer to kinds, not only in the case of kind-level predication but also in the case of object-level predication, as originally proposed in Carlson (1977). Their quantificational force, however, is governed by the same principles that have been established for the variable quantificational force of regular indefinites (see Dayal 2011a and references cited there for background).

Chierchia (1998) extends the kind-based view of English bare plurals to a cross-linguistic theory of noun phrase variation. Within the general perspective of flexible types (Partee 1987), he admits three basic operations for turning an NP with a predicative meaning (type \(<e,t>\)) into an argument (type \(e\) or \(<<e,t>t>\)), nom, iota and \(\exists\):

\[
\begin{align*}
(17a) & \quad \text{iota:} & \lambda P \; t \; P_x, \; \text{if there exists a unique maximal entity in} \; P, \; \text{undefined otherwise.} \\
(b) & \quad \text{nom:} & \text{For any property} \; P \; \text{and world/situation} \; s, \; \cap P = \lambda s \; t P_x, \; \text{if} \; \lambda s \; t P_x \; \text{is in} \; K[\text{IND}] 
\end{align*}
\]
undefined otherwise, where \( P_s \) is the extension of \( P \) in \( s \).

(c) \( \exists \lambda P \lambda Q \exists x [P(x) \& Q(x)] \)

Of these, he considers the first two meaning preserving in the sense that they map a predicate into an entity without introducing quantificational complexity. \( \iota \) picks out the unique maximal entity in the extension of the predicate at the relevant situation, if there is one, and is undefined otherwise. As indicated in section 1, this is the lexical meaning of the definite article in English, but in languages without a definite article it operates as a covert type shift. \( \text{Nom} \), the kind forming operator, is a function from indices to the maximal entity that is in the extension of the predicate at that index – that is, it yields the unique maximal entity that instantiates the kind at the index. \( \text{Nom} \), which is not lexically encoded in English, is a partial function because it is undefined for predicates that do not fit the concept of a kind. It also only operates on plural predicates. The third type-shift \( \exists \) not only turns a predicative expression into an argument, it also introduces \( \exists \) quantification. This is the meaning of the determiner \( a/an \) in English. Since this yields a generalized quantifier, it can interact with other scopal expressions. Unlike the first two operations, \( \exists \) is total. These possibilities are constrained by two principles specific to type shifts:

\begin{align*}
(18a) & \quad \text{Ranking: } \text{nom} > \{\iota, \exists\} \quad \text{to be revised} \\
(b) & \quad \text{Blocking Principle (‘Type Shifting as Last Resort’): } \text{For any type shifting operation } \pi \text{ and any } X: * \pi(X) \text{ if there is a determiner } D \text{ such that for any set } X \text{ in its domain, } D(X) = \pi(X). 
\end{align*}

Finally, the rule of \( \text{DKP} \) mediates between a kind denoting term and a predicate of objects, a repair operation on sorts. Essentially, it existentially binds instantiations of the kind at the relevant index. Since this \( \exists \) is introduced at the point where the sort adjustment is required, it ensures obligatory narrow scope for its operand:

\begin{align*}
(19a) & \quad \text{DKP: If } P \text{ applies to objects and } k \text{ denotes a kind, then } P(k) = \exists x [\lnot k(x) \land P(x)] \\
(b) & \quad \text{PRED (\text{‘\&’})}: \lambda d_s.e. \{\lambda x [x \leq d_s] \text{ if } d_s \text{ is defined, } \lambda x[\text{FALSE}] \text{ otherwise}\}, \text{ where } d_s \text{ Is the plural individual that comprises all of the atomic members of the kind.}
\end{align*}

A simple demonstration of the system can be given for English:

\begin{align*}
(20a) & \quad \text{Dogs have evolved from wolves} \\
(b) & \quad \text{evolve-from}(\text{dogs, wolves}) \\
(21a) & \quad \text{Dogs are barking.} \\
(b) & \quad \exists [\text{are-barking (s)} (\text{dogs(s)(x)})] = \text{DKP} \Rightarrow \exists s x [\lnot \text{dogs(s)(x)} \land \text{are-barking (s)(x)}]
\end{align*}

Since \textit{evolve} in (20a) is a kind level predicate, and the predicates \textit{dogs} and \textit{wolves} have the requisite intensionality, \textit{nom} turns them into arguments which can be fed into the verb meaning directly. Since \textit{bark} in (21a) is an object-level predicate, it cannot hold of the kind, only of its instantiations and \( \text{DKP} \) comes into play. The truth conditions associated with (21b) are identical
to those of a corresponding statement with an indefinite but there is a difference that shows up in scopal contexts. Even when a bare plural outscopes negation, for example, it ends up having narrow scope:

\[(22a)\] Dogs are not barking
\[(b)\] \[dogs_i \ [not \ [t_i \ are \ barking]] \ ]
\[(c)\] \(\lambda x_i \ [[not \ [t_i \ are \ barking]]] \ (\ [\cap \ dogs] \ )\]
\[\Rightarrow \ \text{are-barking} \ (s) \ (\cap \ dogs) = \text{DKP} \Rightarrow \ \exists x [\cup \cap \ dogs(s)(x) \ \land \ \text{are-barking}(s)(x)]\]

Since the bare plural is individual denoting (type \(<s, e>\), it gets lowered into the argument position of the negative predicate. When DKP adjusts the mismatch between \textit{barking} and \(\cap \ dogs\), \(\exists\) enters into the derivation, necessarily below negation. The regular indefinite instead is a generalized quantifier, which means that it enters into an operator-variable relation with its trace and can therefore have scope over negation. Appealing to reference to kinds for bare plurals and to a generalized quantifier meaning for indefinites thus yields radically different scopal properties for them. This difference, we saw, is particularly striking in the case of differentiated scope readings in which bare plurals have readings unavailable to indefinites.\(^{11}\)

The ranking proposed by Chierchia requires revision since it does not capture the facts that he wants to capture, namely that bare nominals in languages without articles can be definite as well as indefinite. The ranking in (18a) predicts that both \textit{iota} and \(\exists\) would be outranked by \textit{nom}, making both definite and indefinite readings unavailable. We have argued against the (in)definiteness of bare nominals in Hindi. As such, the ranking of type shifts that is needed is the one in (23) from Dayal (1999, 2004):

\[(23)\] \(\{\text{nom, iota}\} > \exists\)

The ambiguity of Hindi bare nominals as definites and kinds, but not indefinites, now follows straightforwardly. The only indefinite readings possible are derivative on their kind reading or due to properties specific to certain syntactic positions or constructions. Hindi and English bare plurals differ, of course, in the possibility of the definite reading because \textit{iota} is available in Hindi as a covert type shift but is lexically blocked in English by the definite determiner .

2.2. **Singular vs. Plural Bare Nominals**

The revised ranking of type shifts explains the properties of Hindi bare plurals. We now need to account for differences between Hindi bare plurals and Hindi bare singulars. Let us focus on the differentiated scope reading, illustrated in (12b) and repeated below:

\[11\] The unavailability of differentiated scope readings for indefinites is one of the strongest arguments for adopting the neo-Carlsonian approach over an approach that treats bare plurals as ambiguous between kinds and indefinites (see Dayal 2011a for a survey).
To recap, the bare plural names a kind and is of type \( \langle s,e \rangle \). It can therefore function as a direct argument of the verb. If the verb is not a kind-level predicate, as in this case, there is a sort mismatch which DKP repairs. We get the narrowest scope reading in cases like these where the adverbial operator takes scope immediately above the verb and below the position at which a generalized quantifier is interpreted. Thus the existential associated with the bare plural kind term takes scope at a position unavailable to the indefinite. But if bare singulars are also kind terms but do not have DKP-based indefinite readings, an explanation is called for.

The account of singular kind terms proposed in Dayal (2004) takes them to be semantically plural but grammatically atomic, in analogy with collective nouns like committee or team. Briefly, it is proposed that nom is not responsible for singular kind terms. Rather, they are formed by the application of iota to a property of taxonomic/sub kinds. In languages in which iota is lexicalized, we get the definite singular generic, in languages in which it operates as a covert type-shift we get bare singular kinds, as in the singular variant of Hindi (12a):

(25a) The lion is an endangered species
(b) endangered-species(\( \text{ix } [\text{LION}(x)] \))

The relevant point here is that such terms do not allow access to their instantiations for purposes of predication. Any indefinite reading of bare singulars, then, must be based on its predicative meaning interacting with construction-specific factors. For example, in direct object position a bare singular can be treated as a predicative term of type \( \langle e,t \rangle \), which feeds into the meaning of an incorporating verb and ends up having existential force. Thus a non case-marked bare singular in object position, but not in subject position, can show differentiated scope. I refer the reader to Dayal (2004) for the full picture on singular kind terms and to Dayal (2011b) for their indefinite reading in direct object position. Similar reasoning applies to the existential construction in (14b).

Let me end the discussion of kind terms by making two points. The first is that Hindi bare singulars and English definite singulars can both function as kind terms but Hindi bare singulars can also function as predicates. This explains why bare singulars seem to have a wider range of readings than definite singular generics. It still leaves open why it is possible, to whatever extent, to use at least some bare singulars in presentational contexts. Note that we cannot appeal to a generalized quantifier interpretation for those cases as it would incorrectly predict that they would behave like indefinites more generally. The second point is that in languages in which plural kind terms must occur with a definite determiner, such as the Romance languages, kind terms do not have DKP based indefinite readings. This suggests that the presuppositions associated with a lexical exponent of nom or iota cannot be overridden by DKP while those that rely on covert type shifts can. Due to constraints of space, however, I must leave further discussion of these issues for another occasion.
2.3. Syntactic Issues

The neo-Carlsonian explanation for the Hindi facts was couched in a view of bare nominals as syntactic NPs. But it is also conceivable that such nominals are syntactic DPs. Let us see whether making this move has consequences for the theory discussed above, where a bare nominal with a basic predicative meaning must resort to covert type shift to achieve semantic status as an argument. As such, the principles of Blocking and Ranking, crucial to a cross-linguistically valid theory of meaning, are defined in terms relevant to covert type shifts. However, the issues at the heart of the account would not be affected by a shift from an NP analysis to a DP analysis for bare nominals. Let us see why by considering two possible implementations of such an approach:

26a. \([\text{DP} \emptyset [\text{NP} \text{dogs}]]\)

b. \([\text{DP} \text{dogs}; [\text{NP} t]]\)

There is a null determiner in (26a), while the N head moves from inside NP to D in (26b). Such proposals have been made in the literature and I will briefly mention a few instances in the course of the discussion, without any claim to giving an exhaustive list. The goal here is to demonstrate the relationship between these syntactic options and the issue of indefinite vs. definite interpretation.

Let us assume, then, that everything we have said about Hindi bare NPs holds true but that the structure of the bare nominals involves a null determiner. Based on what we have seen in section 1, we would have to say that this null determiner encodes two type shifts, \(\text{iota}\) and \(\text{nom}\). Crucially, it does not encode the \(\exists\) type shift. If we wanted to derive this in a principled way and to make cross-linguistic predictions we would have to say that null determiners are subject to the same \(\text{Ranking}\) that was proposed for covert type shifts, namely \(\{\text{nom}, \text{iota}\} > \exists\). Moreover, we would have to say that the \(\text{Blocking Principle}\) applies to null determiners in the same way that Chierchia claimed for covert type shifts. Only then could we explain why bare plurals in languages like English do not have definite readings while bare nominals in languages like Hindi do. In other words, positing a null determiner does not provide any substantive advantage as far as interpretation is concerned, as Carlson had already pointed out in (1977). In this case, it simply changes the terms of the discussion from talking about covert type shifts to covert determiners but leaves untouched the essence of the claims.

One final consideration has to do with a possible requirement that the null determiner be syntactically licensed. While clearly some kind of licensing requirement holds for bare plurals in Italian (Longobardi 1994, Chierchia 1998), which are generally acceptable in direct object position but not in subject position, it does not transfer over to English bare plurals or Hindi bare plurals which do not show any subject-object asymmetries. And even for Italian the claim of a null determiner is not enough to settle the semantic question. Chierchia has argued that the null determiner encodes \(\text{nom}\) and under that view it follows that the indefinite readings associated with Italian bare plurals can only have narrow scope readings. Longobardi, among others, have argued that Italian bare plurals unlike English bare plurals are not kind terms. But on that view, there is no explanation for the fact that they have obligatorily narrow scope readings. It would
seem then that if they are not kind terms and they are restricted to object position, they may well be cases of NPs pseudo-incorporated by the verb.\textsuperscript{12}

Now let us turn to the second option, given in (26b). Note that such structures typically have an effect on word order, as shown for Italian by Longobardi (1994). Such structures have also been claimed for articleless languages such as Bangla (Bhattacharya 1999). The base order of Bangla numeral constructions is the one given in (27a) and the derived order in (27b), with N→D being semantically triggered by definiteness/specificity:

\begin{align*}
27a. &\quad [\text{NumP} \text{du} [\text{CL-P} \text{To} [\text{NP boii}]]] \\
&\quad \quad \text{two} \quad \text{CL} \quad \text{book} \\
&\quad \quad \text{“Two books”} \\
\text{b.} &\quad [\text{DP boii} [\text{NumP} \text{du} [\text{CL-P} \text{To} t_1]]] \\
&\quad \quad \text{Book} \quad \text{two} \quad \text{CL} \\
&\quad \quad \text{“The two books”}
\end{align*}

I have argued in Dayal (2012, 2014) that (27a) has an indefinite reading, specific as well as non-specific, while (27b) has only a definite reading. Setting aside several important details, the relevant point from the present perspective is that N→D movement is not enough to provide an account of available meanings. \textit{Ranking} has to be included in the theory to predict the definiteness of such structures and the absence of indefiniteness, specific and non-specific.

The more general point is that positing a DP structure with a null D or a DP structure with N→D movement does not replace the task of semantic explanations. Absent a fully worked out compositional account, there is no particular advantage in making the move from the theory given in 2.2 which is based on bare nominals as NPs undergoing type shift. I should clarify, however, that there is no inherent incompatibility between the neo-Carlsonian account and a DP based account. I simply want to highlight the need for an independently defensible semantics to go with whatever syntactic option one proposes.

3. \textbf{Further Implications}

In this section we take up two areas where the findings for Hindi proposed in sections 1 and 2 have significant implications. We first consider the possibility of cross-linguistic variation by exploring whether the theory allows for article-less languages in which bare nominals are indeed ambiguous between definite and indefinite readings. We then consider the well-known fact that L2 learners of articulated languages coming from an article-less L1 face problems with the determiner system. Here we look at recent experimental work in this area and draw out the implications of the proposal forwarded here that bare nominals in article-less languages are not truly ambiguous between definite and indefinite.

\textsuperscript{12} Thanks to Ivano Caponigro and Gennaro Chierchia for discussions related to this issue. Note that for Chierchia’s kind-based account of this to work, he has to exempt the null determiner from competition with the overt plural definite determiner or give a somewhat different semantics to the definite determiner. He inclines to the latter.
3.1. (In)definiteness in Other Article-less Languages

Let us start by briefly considering the extent to which the insights gleaned from our study of Hindi bare nominals can inform investigation into the semantic properties of bare nominals in other article-less languages. The first issue to settle for any language is whether it has definite and indefinite lexical determiners. We know that definite determiners develop from demonstratives and indefinite determiners from the numeral one. Demonstratives and numerals are attested in most, perhaps all languages, so we need diagnostics to identify lexical determiners. We have already seen diagnostics for identifying definite determiners, namely anaphoricity and homogeneity (cf. section 1.2). Note though that anaphoricity may also be possible with demonstratives. It is, therefore, important to probe the uniqueness inherent to definites in a different way, using the homogeneity test. One possibility is to take a context in which there are several children and see whether a singular bare nominal can be used felicitously. If it can, perhaps with a pointing gesture, we have evidence that the determiner is a demonstrative not a definite:

(28a) That child is standing and that child is sitting.

b. #The child is standing and the child is sitting.

Related to this is the fact that demonstratives have a non-uniqueness implicature (Robinson 2005). If there is only one student in the context, it is a bit odd to say that child is sitting.

There are two diagnostics for separating indefinite determiners from the numeral, discussed in Chierchia (1998). Indefinite determiners lend themselves to generic statements and are able to take scope under negation:

(29a) A student works hard.

(b) John didn’t buy a book.

The Hindi translation of (29a) with ek vidyaarthii ‘one student’, for example, lacks the reading that working hard is the property of students generally. And the Hindi translation of (29b) with ek kitaab ‘one book’ does not have a neutral narrow scope reading, conveying instead a wide scope reading for the indefinite or an emphatic narrow scope reading. In this respect it is similar to John didn’t read one book. We can say with some confidence, then, that any language that aligns with Hindi in these respects lacks both the definite and the indefinite articles.

It is worth noting, though, that while the presence or absence of a definite determiner makes a clear difference in the possibility of definite readings for bare nominals, the presence or absence of an indefinite determiner may not have an impact on indefinite readings. Recall that we have proposed two constraints on covert type shifts: blocking and ranking (cf. 18). Although blocking is not without its complexities (see Dayal 2013), it is still a good predictor of the availability of definite readings for bare nominals. But if nom and iota outrank ∃, it is predicted that bare nominals will not have a genuine indefinite interpretation regardless of the presence or absence of an indefinite determiner in the language. Therefore, regardless of whether the
language has an indefinite determiner or not the theory as currently stated will predict the absence of genuine indefinite readings for bare nominals in articleless languages.

Of course, it is entirely possible that ranking is subject to cross-linguistic variation, and if so, bare nominals in a language that does not follow this principle would have all three readings: kind, definite and indefinite. While this is a possibility, one of the lessons from the case study of Hindi in section 1 is the importance of going beyond intuitions regarding the novelty-familiarity of the referent in determining the status of a bare nominal as indefinite. Tests probing specificity and scope were shown to be significant in this connection. If a bare nominal seems to take obligatory narrow scope, even in contexts where an indefinite would not, we can be relatively sure that it is not a bona fide indefinite. The source of the perceived indefiniteness would then have to be constructional or derivative on a kind level meaning. Direct objects, we saw, can be targets of incorporation and/or complex predicate formation. Similarly, subjects may involve topicalization and therefore lean towards a definite interpretation. It is thus important to check for indefinite readings in a variety of syntactic positions. The indirect object position, which has neither of these biases, proved to be revealing in the case of Hindi bare singulars.

As mentioned above, kind terms are one source for narrow scope indefinite readings of bare nominals. It may be worth highlighting that their status as kind terms needs to be established on the basis of kind-level predicates rather than on the basis of generic statements. As observed by Carlson (1977), this test separates English indefinites, which are compatible with generic statements, from bare plurals which are genuine kind terms. (30b) is only acceptable under a taxonomic reading, referring to a subtype of dinosaur:

\[(30a) \text{ Dinosaurs are extinct. } \]
\[(30b) \text{ * A dinosaur is extinct. } \]

Note in this connection that Hindi and English both mark number inside the nominal and that singular and plural kind terms differ in two respects. Languages with articles typically disallow bare singulars and do not typically admit indefinite readings for singular kind terms. In languages that do not make a similar distinction but allow bare nominals to refer to kinds, bare nominals typically behave like bare plurals rather than bare singulars. This holds for languages like Chinese that require the obligatory use of classifiers for counting as well as for languages like Haitian Creole that do not. So far I have not come across languages in which bare nominals are bona fide indefinites, displaying the range of readings associated with indefinites.

### 3.2. From Article-less L1 to Articled L2

The statement that adult learners of a language with articles never quite master the system if their L1 lacks articles is almost a truism. This observation has recently been validated with systematic

---

13 The crucial contrasts were demonstrated in (12b). Unfortunately, these cases have been generally neglected in the literature leading to the erroneous view that bare plurals are simply a sub-type of indefinites.

14 The definite singular generic lends itself to representative object interpretations, as in *We photographed the lion on our trip* or *The rat reached Australia in 1770* (see Dayal 2011a for a survey).
tests using modern methods of statistical experimentation and analysis. In this final section I
would like to briefly consider this issue from the perspective of the view espoused in this paper
about bare nominals in article-less languages. In particular, it makes claims about the types of
errors adult learners are likely to make: errors of misuse versus errors of omission. To get a sense
of what is at issue, consider the following paradigm:

31a. There were ten children in the room.  
   ___ girl was playing…  Hindi L1  English L2
   ek ‘one/*Ø  no error
b. Our rules say that
   if ___ student registers for a class, it will
   not be canceled.  koi/*Ø  no error
   koii/*Ø  no error
   Ø  omission
c. A child came in.
   ___ child sat down.

The paradigm is modeled on Ionin (2003), where subjects are given a forced-choice tasks in
which they are asked to choose between THE, A or Ø to fill in the blank. The target for the
above, would be A, A and The. The prediction that the current theory of Hindi bare nominals
makes is that adult learners would not make errors in (31a) and (31b) since transfer from their L1
would suggest the need for a lexical exponent which is the closest correspondent to ek or koii. In
the case of (31c) the prediction would be that there would be errors of omission as the bare
nominal would be fully accepted in their L1.

As would be obvious, the paradigm in (31) targets contexts where the L1 situation is
unambiguous. No Hindi speaker would use a bare nominal for the first two cases, and no Hindi
speaker would use the numeral or the indefinite determiner in the third case. We know, of
course, that had we chosen examples which targeted the direct object position, the L1 situation
could be construed as favoring the definite or the indefinite. This is particularly so in the case of
written tasks that do not have the benefit of prosody to disambiguate. One general question this
raises, of course, is whether interference from L1 is sensitive to specific constructions or is more
global. Is it the case that if a bare nominal can be construed as definite or indefinite in some
constructions, the ambiguity becomes a factor in L2 learning regardless of the particular
construction. Or is it the case that interference from L1 will be context dependent, as shown
above?

Although there are fairly extensive studies on L2 acquisition of determiner systems at this
point (Ionin 2003, Ko et al 2010, Schönenberger 2014, among others), they do not target this
particular question. The basic premise of most such studies is that bare nominals in article-less
languages are ambiguous. In fact, the following quote from Schoenenberger (2014: 84) about the
experimental design being reported on is telling, “to ensure that word order would not influence
article choice, all the test items contained transitive verbs and article choice always concerned

15 The L1 in the studies I have looked at are not Hindi. Although I do not want to make claims here for other
languages, as I have already indicated, I have not found bare nominals in the L1 referred to in these studies
(Russian, Korean, for example) to have bona fide indefinite readings.
nominals in the object position. These nominals were singular count nouns, which always require an article.” Now, as we have seen, the direct object position is precisely the context where bare singular count nouns in the L1 article-less languages are prone to ambiguity. It would be worthwhile for further studies in this domain to control for syntactic position and interpretation as a way to probe more deeply and compare the results with those for direct object position. A second comment, also from Schoenenberger (2014: 99-100), is interesting in the context of our findings in sections 1 and 2, “article omission is significantly higher with definites than with indefinites.” This is exactly what one would expect if bare nominals in L1 are definite rather than indefinite.

4. Conclusion

In this paper I have collected together various tests that I have used in the past to determine the semantic properties of bare nominals. In the process, I have emphasized the need for precision and clarity in making claims about their interpretation. Although much more is known about the semantics of bare nominals today than even twenty years ago, we still need to flesh out an empirically adequate picture of noun phrase interpretation in order to refine our current understanding of cross-linguistic variation. Shifting our attention from languages with overt exponents of (in)definiteness to those without articles can help us gain a deeper understanding of what constitutes definiteness and indefiniteness. And making headway in this domain can also shape the way in which we probe issues related to the acquisition of the determiner systems by bilinguals and add to our understanding of how (in)definiteness is encoded in the human brain.

References


