(In)definiteness in Article-less Languages
And Article-use in New Englishes

Veneeta Dayal
Yale University

Randolf Quirk Fellow Workshop 3
Queen Mary University of London
May 22, 2024
A Little Housekeeping – Workshops 1 & 2

- The anti-uniqueness presupposition of demonstratives, is taken from Dayal & Jiang (2021) – as mentioned in the handouts from Workshop 1 and Workshop 2, but it failed to mention that Dayal & Jiang refer to Robinson (2005) as the source of their claim.
- I have been suggesting that most of the research on demonstratives within formal semantics has focused on its indexicality. But I should point out that Robinson (2005), obviously, has taken anti-uniqueness into account. Furthermore, Wolter (2006) also has an account of demonstratives that traverses much of the same ground.
- I need to go back and look at Wolter’s work again to see if it can capture the contrasts I have focused on, and if not, whether it can be supplemented to do so (specially in light of suggestions by Matthew Husband and Tom Stephens that the anti-uniqueness of demonstratives may be derivable via competition with the uniqueness of definites, in approaches that posit widened domains for – details to be worked out.


https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8ac80e9f8bddd125b72c476336d1f4b3039b611b

The Open Handbook of (In)definiteness:  
A Hitchhiker’s Guide to Interpreting Bare Arguments  
V. Dayal (ed), to appear, Open Handbooks in Linguistics, MIT Press

TOC

Foreword: Gennaro Chierchia

PART I: From (silent) morpho-syntax to meaning

Chapter 1: From (silent) morpho-syntax to meaning: An Introduction Veneeta Dayal
Chapter 2: A guide to some semantic concepts Veneeta Dayal
Chapter 3: The (in)definiteness questionnaire Veneeta Dayal

PART II: Case Studies

Chapter 4: (In)definiteness in Cabo-verdean Creole – a Case Study Marlyse Baptista & Veneeta Dayal
Chapter 5: (In)definiteness in Cuzco Quechua – a Case Study Liliana Sanchez, Janett Vengoa & Veneeta Dayal
Chapter 6: (In)definiteness in Hiaki – a Case Study Heidi Harley, & Veneeta Dayal
Chapter 7: (In)definiteness in Indonesian – a Case Study Daniel Kaufman, Gita Martohardjono & Veneeta Dayal
Chapter 8: (In)definiteness in Korean – a Case Study Sea-Hee Choi, James Yoon & Veneeta Dayal
Chapter 9: (In)definiteness in Russian – a Case Study Anita Soloveva, Maria Polinsky & Veneeta Dayal
Chapter 10: (In)definiteness in Xhosa – a Case Study Vicki Carstens, Loyiso Mletshe & Veneeta Dayal

PART III: Lessons Learned

Chapter 11: Case Studies on (In)definiteness: Empirical Generalizations Veneeta Dayal
Chapter 12: Case Studies on (In)definiteness: Implications for Theory Veneeta Dayal

PART IV: (In)definiteness Beyond the (In)definiteness Questionnaire

Chapter 13: Cabo Verdean Creole and Mandinka DP Domain: A Comparative Analysis Marlyse Baptista and Ousmane Cisse
Chapter 14: Is the Cuzco Quechua plural marker –kuna optional? Liliana Sanchez & Janett Vengoa
Chapter 15: Probing (In)definiteness in an Endangered Language Heidi Harley
Chapter 16: Definiteness marking in Austronesian: an overview Daniel Kaufman
Chapter 17: Factors that influence the use of plural marker tul in Korean Sea-hee Choi and James Yoon
Chapter 18: Split Scope Definites in Russian Anita Soloveva
Chapter 19: The Structure of the Xhosa Noun Phrase Vicki Carstens

Afterword: Claire Bowern
Goals

- To evaluate the claim “bare arguments in languages without articles are ambiguous between definite and indefinite”.
- There are two ways to interpret this statement: as an existential statement or as a universal statement.

The existential claim: Bare arguments in article-less languages can be interpreted as definite and as indefinite.

\[ \text{The} \cap \text{A} \]

The claim would be true as long as a bare argument could be used in some contexts where ‘the’ can be used and some context where ‘a’ can be used. This is undoubtedly true.

The universal claim: Bare arguments are interpreted as definites (hence useable wherever ‘the’ is usable and wherever ‘a’ is usable).

This claim is be true iff the full logical space in the diagram above can be covered by bare arguments in article-less languages.
Löbner (1985: 320): “as for languages which do not have a definite article, it is plausible to assume that they just do not explicitly express the way nouns are to be interpreted.”

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”.

This claim is not true for all article-less languages

Results of the Case Studies: of the 7 languages investigated

4 were unequivocally definite only (CVC, CQ, Indonesian, Russian)
1 was unequivocally definite and indefinite (Xhosa).
Novelty-Familiarity: the articles signal novelty/familiarity

Conversation over the phone between two people:

1a. Speaker A: What’s that noise?

   b. Speaker B: Oh, a dog is barking outside my door.  
   c. Oh, the dog is barking outside my door.

   No prior knowledge of a dog on the part of A; no uniqueness  
   Familiar to both; uniqueness  
   (# if B has two pet dogs, for example)

Note: If both participants have access to the situation, it is possible to refer to a novel entity with the indefinite or the definite (possibly via accommodation).

Word Order can signal (in)definiteness in article-less languages like Hindi and Russian:

2a. maidaan meN kuttaa hai  
    park in dog is  
    “There’s a dog in the park”  

   b. kuttaa maidaan meN hai  
      dog park in is  
      “The dog is in the park.”

3a. na trave sidit sobaka  
    on grass sit dog  
    “There’s a dog sitting on the grass”.

   b. Sobaka sidit na trave  
      Dog sit on grass  
      “The dog is sitting on the grass.”
I. Russian and Hindi Bare Arguments

So even in languages that I have classified as belonging in the set of languages where bare nouns are unequivocally definite on the basis of a full set of diagnostics, bare arguments are able to introduce new discourse referents.

4a. $[[D_{-DEF}]] = \lambda P \lambda Q \exists x [P(x) \land Q(x)]$

b. $[[D_{+DEF}]] = \lambda P \iota x [P(x) \land x=y]$ to be refined

5a. $\exists x [\text{dog}(x) \land \text{in-the-park/on-the-grass}(x)]$ Hindi/Russian post-locative bare N

b. in-the-park/on-the-grass($\iota x (\text{dog}(x) \land x=y)$) Hindi/Russian pre-locative bare N

The proposal in (4) is for the two basic articles, it can be extended to other expressions such as pronouns, demonstratives, numerals, *some* etc. There are no determiners that can be used anaphorically and introduce discourse referents:

6a. #Exactly one student$_i$ has registered for this course. A/One/Some student$_i$ (who registered) will be disappointed if we cancel.

b. Exactly one student$_i$ has registered for this course. She$_i$/The student (who registered$_i$)/That student$_i$ will be disappointed if we cancel.

7a. #Once upon a time, there was the/that man who had the/that daughter.

b. Once upon a time, there was a man who had a/one daughter.

Novelty-Familiarity is a feature of (overt) determiners – it doesn’t apply to languages that don’t have determiners. Hindi and Russian bare arguments can therefore introduce DRs and be anaphorically linked to previously introduced DRs.
I. Russian and Hindi Bare Arguments

(Non-)Uniqueness and (In)definiteness

Imagine calling an animal shelter and asking what two of your interns are doing and getting (56b) in response:

8a. A dog is running around in the park and a dog is sleeping near the sandbox.
   b. Mary and Sue are both busy. Mary is walking a dog and Sue is bathing a dog. |dog| ≥ 1
   c. # Mary and Sue are both busy. Mary is walking the dog and Sue is bathing the dog. |dog| = 1

2a. maidaan meN kuttaa hai  Hindi  # if 2 equally salient dogs
   park in dog is “There’s a dog in the park”
3a. na trave sidit sobaka  Russian  (Anita Soloveva, pc)
   on grass sit dog  “There’s a dog sitting on the grass”.  # if 2 equally salient dogs

Dayal (2004): non-familiar definites; do not presuppose existence, but convey uniqueness/maximality

Downward Entailing Contexts and ‘indefinite’ Readings
Negation and antecedents of conditional – the uniqueness requirement projects out.

Judgments gratefully acknowledged: Rajesh Bhatt for the Hindi data,
Elena Koulidobrova and Fedor Golosov for the Russian data.
I. Russian and Hindi Bare Arguments

In (9)-(10), with negation added, the sentences become significantly degraded Interestingly, there is no problem with the corresponding sentences with bare plurals:

9a. ?? maidaan meN kutta nahiiN hai
    Park in dog not is Intended: there isn’t a dog in the park. unavailable: ¬∃

   b. maidaan meN kutte nahiiN haiN
      park in dogs not are There aren’t any dogs in the park. available: ¬∃

10a. ?? na trave ne sidit sobaka
      On grass not sit dog Intended: “There isn’t a dog sitting on the grass” unavailable: ¬∃

   b. na trave ne sidiyat sobaki
      on grass not sit dogs “There aren’t dogs sitting on the grass” available: ¬∃
I.Russian and Hindi Bare Arguments

A similar contrast between singular and plural surfaces in (11) and (12). In both languages, the bare singular in the antecedent of the conditional suggests that there is only one child or student in the context and this comes in the way of possible narrow as well as wide scope $\exists$ readings. To get the narrow scope indefinite reading *if any child/student comes, call me* an overt determiner (*koii* ‘some’ in Hindi, *kakoy-to* ‘some’ in Russian) is needed. To get the wide scope indefinite reading *if a particular child/student comes, call me*, another overt determiner is needed (*ek* ‘one’ in Hindi, *odin* ‘one’ in Russian):

11a. ? agar baccaa aaye, mujhe phauran bataanaa
if child come me-DAT immediately tell
Intended: “If any child comes in, call me immediately.”

b. agar bacce aayeN, mujhe phauran bataanaa
if children come me-DAT immediately tell
“If any children come in, call me immediately.”

12a. ? yesli student budet zvonit’, dayte mne znat’
If student will call let me know
Intended: “If a student comes in, call me immediately.”

b. yesli studenty budut zvonit’, dayte mne znat’
If students will call let me know
“If any students come in, call me immediately.”
I. Russian and Hindi Bare Arguments

Proposal:
Hindi & Russian bare singulars are not simple indefinites – we have seen they do not allow (13b) and (13c); rather they have a singleton set requirement on N as in (14); plurals do not have any restrictions on size (15).

13a. \( \exists x \) \([\text{dog}(x) \land \text{in-the-park/on-the-grass}(x)]\)
   
   b. \( \neg \exists x \) \([\text{dog}(x) \land \text{in-the-park/on-the-grass}(x)]\)
   
   c. \( \exists x[\text{student}(x) \land \text{come-in}(x)] \rightarrow [\text{call-me}] \)

14a. \( \exists x \) \([|\text{dog}| = 1 \land \text{dog}(x) \land \text{in-the-park/on-the-grass}(x)]\)
   
   b. \( \neg \exists x \) \([|\text{dog}| = 1 \land \text{dog}(x) \land \text{in-the-park/on-the-grass}(x)]\)
   
   c. \( \exists x [|\text{student}| = 1 \land \text{student}(x) \land \text{come-in}(x)] \rightarrow [\text{call-me}] \)

15a. \( \exists x \) \([|\text{dogs}| \geq 0 \land \text{dogs}(x) \land \text{in-the-park/on-the-grass}(x)]\)
   
   b. \( \neg \exists x \) \([|\text{dogs}| \geq 0 \land \text{dogs}(x) \land \text{in-the-park/on-the-grass}(x)]\)
   
   c. \( \exists x [|\text{students}| \geq 0 \land \text{students}(x) \land \text{come-in}(x)] \rightarrow [\text{call-me}] \)

(14b) is something like: \((there can be only one dog and) it turns out that no dog is in the park.\)

(14c) is something like: \((there can be only one student and) if any student comes, call me.\)

(15b) is something like: \((there may be none or any number of dogs and) it turns out that no dogs are in the park.\)

(15c) is something like: \((there may be none or any number of N and) if any students come, call me.\)

Independent Motivation: In English the singular is only possible with the numeral *one* but the plural is much less restricted:


zero books, 1.5 children, 2 dogs etc. (*one books, blocked by one book)

(Krifka 19xx)
II. Article Use in New Englishes/South Asian English(es)

“For [Indian English] indubitably now is a bona fide dialect of English (not some other dialect, spoken with lots of mistakes), and it deserves to be provided with a comprehensive descriptive grammar.”

Dixon 1991:447

“Substrate languages are a much stronger influence than universals on contact outcomes. However, the new English system is rarely an exact copy of the substrate. New Englishes often develop “third grammars” that derive from L1-driven inferences or discourse-based solutions to the challenge of mapping L1 meanings to surface forms in English”.

Figure 1. Kachru’s Concentric Circles of World Englishes

Figure 2. Modiano’s ‘Centripedic Circles of English as an International Language’ (from Modiano 1999a)
II. Article Use in New Englishes/South Asian English(es)

“The greater entrenchment of some substrate-based features than others is accounted for by appealing to the additional factor of input demand. As IndE and SgE both involved a historical reduction of British English input, SLA models prove very relevant and remind us that language acquisition is at the heart of dialect birth in postcolonial settings.”

Sharma 2023:7: From Deficit to Dialect

“There is no such thing as a single variety of Pakistani English just as there is no single variety of Indian English. Both are convenient labels for several South Asian varieties of English. In Pakistan the languages of borrowing are Punjabi, shared with the Indian Punhab; Urdu, shared in the form of spoken Hindi with India; Sindhi, again shared with migrants from Sindh to India, and Pashto, which, however, is shared with Afghanistan not India.”

Rahman 2020: 283
II. Article Use in New Englishes/South Asian English(es)

**Non Standard Article Use (NSAU):** a ubiquitous feature of SA English (and New Englishes more generally)

“To explore the reorganization of English article use in IndE, we need a basic understanding of how articles are used in Indian languages. As Hindi is the primary substrate for the IndE speakers analyzed here and is representative of article systems found across Indian languages, Hindi is used as the sample L1/substrate here.

The article systems of all Indo-Aryan and Dravidian languages spoken by the current sample of speakers [24 speakers]mark a specific/nonspecific contrast, but not a definite/indefinite contrast. All of the languages optionally use their form for the numeral *one* to indicate specific indefinite meaning (Sridhar 1990, Cardona 1965, Schiffman 1999). And none of the languages have a definite article, instead, definiteness is marked by word order, case marking, and prosody.”

Sharma 2023: 80-81
II. Article Use in New Englishes/South Asian English(es)

```
<table>
<thead>
<tr>
<th>Definite</th>
<th>Indefinite</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>ek (&quot;one&quot;)</td>
</tr>
<tr>
<td>-</td>
<td>(including generic)</td>
</tr>
</tbody>
</table>

Figure 2. Hindi article system (for singular count nouns).

<table>
<thead>
<tr>
<th>Definite</th>
<th>Indefinite</th>
</tr>
</thead>
<tbody>
<tr>
<td>the, a (including generic)</td>
<td>a</td>
</tr>
</tbody>
</table>

Figure 1. English article system (for singular count nouns).

Sharma 2005b: 539/Sharma 2023: 81

“The simplified diagrams show that this system is the inverse of the English article system, which marks a definite/indefinite contrast but not a specific/non-specific one.”

**BUT** the grammar of *Hindi does not have a specific/nonspecific contrast!*

The nonspecific reading needs *koii* ‘some’

13a. *agar bacca aaye, mujhe phauran bataanaa*  
if *child* comes me-DAT immediately tell  
*presupposition: |child| = 1*

a’. *agar vidyarthii aaye, mujhe phauran bataanaa*  
if *student* comes me-DAT immediately tell  
*presupposition: |student| = 1*

see Dayal 2017 for more
II. Article Use in New Englishes/South Asian English(es)

Non Standard Article Use (NSAU)

Sharma 2023

Hypothesis 2: 
\([+\text{spec}, -\text{def}] \rightarrow \text{overtly marked ‘one’} \)  
\([-\text{spec}, -\text{def}] \rightarrow \emptyset \) (but \textit{koii!!})

\([+\text{spec}, +\text{def}] \rightarrow \emptyset \)  
\([-\text{spec}, +\text{def}] \rightarrow \emptyset \)

Hypothesis 3: 
\([+\text{SR}, +\text{HK}] \rightarrow \text{overtly definite} \)  
\([+\text{SR}, -\text{HK}] \rightarrow \text{overt indefinite} \) (but \textit{koii!!})  
\([-\text{SR}, -\text{HK}] \rightarrow \emptyset \)  
\([-\text{SR}, +\text{HK}] \rightarrow \emptyset \)

SR: Specific Reference  
HK: Hearer Known

Neither of these hypotheses is strongly confirmed aside from one part of Hypothesis 2 (transfer of the specific/nonspecific distinctions in Indian Languages to IndE in the omission of the indefinite article ‘a’).

Sharma 2023: 83-84
### II. Article Use in New Englishes/South Asian English(es)

**Table 4.2. Spec Ind (ek) Nonspec Ind (θ) Definite (θ)**

<table>
<thead>
<tr>
<th>Group</th>
<th>Spec Ind (ek)</th>
<th>Nonspec Ind (θ)</th>
<th>Definite (θ)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n  %omitted</td>
<td>n  %omitted</td>
<td>n  %omitted</td>
<td></td>
</tr>
<tr>
<td>Group 1</td>
<td>29  48%</td>
<td>442 81%</td>
<td>83  46%</td>
<td>least proficient</td>
</tr>
<tr>
<td>Group 2</td>
<td>129 31%</td>
<td>206 57%</td>
<td>343 38%</td>
<td>more proficient</td>
</tr>
<tr>
<td>Group 3</td>
<td>116 16%</td>
<td>132 20%</td>
<td>209 18%</td>
<td>most proficient</td>
</tr>
<tr>
<td>Total</td>
<td>274 27%</td>
<td>380 47%</td>
<td>635 33%</td>
<td></td>
</tr>
</tbody>
</table>

- Sharma notes that specific indefinites, which would have an overt article in Indian languages, have significantly higher rates of overt articles (lower rates of omission) than non-specific indefinites. This declines with proficiency.
- It is also predicted that there should be a high rate of omission for ‘the’, similar to the high rate of omission of non-specific ‘a’, due to lack of an overt form in L1s.
- But omission rates for the definite article are more in line with the lower rates of omission for specific ‘a’. Neither exceeds 50%.
- Sharma provides an analysis of her results in terms of combination of Hypothesis 2 plus Hypothesis 4, based on Prince 1981’s Taxonomy of Assumed Familiarity.
II. Article Use in New Englishes/South Asian English(es)

Non Standard Article Use (NSAU)

Table 4.2. Spec Ind (ek) Nonspec Ind (θ) Definite (θ)

<table>
<thead>
<tr>
<th></th>
<th>Spec Ind (ek)</th>
<th>Nonspec Ind (θ)</th>
<th>Definite (θ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>% omission</td>
<td>n</td>
<td>% omission</td>
</tr>
<tr>
<td>Group 1</td>
<td>29 48%</td>
<td>442 81%</td>
<td>83 46%</td>
</tr>
<tr>
<td></td>
<td>lowest proficiency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 2</td>
<td>129 31%</td>
<td>206 57%</td>
<td>343 38%</td>
</tr>
<tr>
<td></td>
<td>more proficient</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group 3</td>
<td>116 16%</td>
<td>132 20%</td>
<td>209 18%</td>
</tr>
<tr>
<td></td>
<td>most proficient</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>274 27%</td>
<td>380 47%</td>
<td>635 33%</td>
</tr>
</tbody>
</table>

Sharma’s results have been replicated in other SA Englishes.

We have noted that the non-specific reading in Hindi-Urdu requires a determiner, ie the bare singular is not capable of non-specific readings.

Then why does the non-specific context → article drop (significantly higher than the definite context)?
II. Article Use in New Englishes/South Asian English(es)

What do we learn from these results about +/- Spec, +/-Def in NSAU in IndE?

A more nuanced analysis of the non-specific cases is needed.

Some of the non-specific cases may match the Size Criterion.

Perhaps some of the non-specific cases involve pseudo-incorporation/complex event formation (marry girl from India, look for job).

Maybe *a* is correlated strictly with *ek* ‘one’, not with having/not having an article-like form.

More attention to incorrect uses of the articles, not just their omission, might also be revealing.