Highlights from the *Hitchhiker's Guide to Bare Nominals*

(In)definiteness and Genericity across Languages

Veneeta Dayal

Yale University

May 11, 2023

The Hitchhiker Project

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

- A questionnaire-based study of bare argument interpretation in seven (unrelated) languages.

The Hitchhiker Project

1st Meeting: March 2019

2nd Meeting: May 2023

Languages Studied	***Collaborators***
Cabo Verdean Creole	Marlyse Baptista
Cusco Quechua	Liliana Sanchez, Janett Vengoa
Indonesian	Daniel Kaufman, Gita Martohardjono
Korean	Sea-hee Choi, James Yoon
Russian	Anita Soloveva, Masha Polinsky
Xhosa	Vicki Carstens, Loyiso Mletshe
Hiaki (in prep)	Heidi Harley
Commentators	Claire Bowern, Gennaro Chierchia
Research Associate	Yagmur Sag

The Hitchhiker Project

The (In)definiteness Questionnaire uses well-established tests from the literature to determine whether a particular lexical item in a familiar or unfamiliar language can be classified as a kind term, a definite or an indefinite.

Section I: The goal is to establish three baseline facts: Does the language have a lexical definite determiner? Does it have a lexical indefinite determiner? Does the nominal system of the language encode number distinctions?

Relevant data from English*:*

1a. #**That**/**the sun** is shining.

b. Mary bought a car. #**That steering wheel**/√**the steering** wheel...
Relationship to uniqueness of the noun complement

2a. A cow/#**one cow** eats grass. *On the generic reading*

b. I didn't buy **a**/#**one book**. *On the neutral narrow scope reading: ¬∃*

3a. **Dogs** lick each other.

b. ***The dog** licks each other. *Even on the generic reading of 'the dog'.*

The Hitchhiker Project

The (In)definiteness Questionnaire contd.

Section II: this section tests whether bare arguments in the language under study can be used to make kind-level as well as generic statements:

Relevant data from English*:*

4a. **The dog**/#**A dog** has evolved from **the wolf**/**#a wolf**.
b. **Dogs** have evolved from **wolves**.

Note: **Indefinites** do not allow reference to kinds.

Section III: This section tests whether bare arguments have two properties associated with definites: compatibility/requirement of uniqueness for its N-set, anaphoric potential. The 1st separates **definite determiners** from **demonstratives**, the 2nd is common to both:

5a. #**That**/**the sun** is shining.
b. Mary bought a car. #**That steering wheel**/√**the steering** wheel...
6. A girl and a boy entered. **The girl/that girl** was talking to **the boy/that boy**.

Note: Neither (5) nor (6) allows **an indefinite** in these positions.

The Hitchhiker Project

The (In)definiteness Questionnaire contd.

Section IV: this section tests the status of bare arguments as indefinites along 6 dimensions, separating regular indefinites from regular definites as well as bare plurals.

- Introduction of discourse referents
- Scope interaction with negation
- Partitive Specificity
- Referential Specificity
- Scopal Specificity
- Differentiated Scope

The Hitchhiker Project

The (In)definiteness Questionnaire contd.

	indef	bare-plural	def
Introduction of discourse referents:	√	√	X

7. Once upon a time, **an old woman/old women/#the old woman** lived in this town.

Partitive Specificity:	√	X	X

8. There were ten students in the room.
A student/#**students**/#**the student** was doing homework,
another student/was playing.

*Differentiated Scope**:***	X	√	X

9a. Miles killed #**a rabbit/#some rabbits/ rabbits**/#**the rabbit** repeatedly.

b. Miles baked #**a cake**/#**some cakes/ cakes**/#**the cake** for two months.

Note: Only **bare plurals** are both kind-denoting and allow differentiated scope readings.

The Hitchhiker Project

My Motivations:

- The belief, based on earlier work, that bare nouns in languages like Hindi don't really behave like indefinites, even if they have some indefinite-like properties. That there is a difference between bare plurals and bare singulars wrt the possible indefinite readings.

- The observation that in cross-linguistic studies of (in)definiteness, there is often insufficient, perhaps no attention paid to reference to kinds.

- A cross-linguistically valid theory of the range of meanings available to bare arguments needs an empirical base that is based on systematic diagnostics for determining **kind-reference**, **definiteness** and **indefiniteness.**

- The Hitchhiker's Guide is a **theoretically informed** but **theory neutral** investigation of bare noun interpretation; meant to be of use to fieldworkers, syntacticians, semanticists or anyone who wonders whether a language has or doesn't have determiners, or what it means for a grammar to not have determiners.

The Hitchhiker Project

Background Assumptions

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 express the way nouns are to be interpreted."

Heim (2011): "in languages without definiteness marking, the relevant "ambigous" 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".

Maximize Presupposition:

$[\![\text{the}]\!] = \mathbf{\lambda P}\ \lambda Q\colon |\mathbf{P}| = \mathbf{1}.\ \exists x\ [P(x) \wedge Q(x)]$ or $\mathbf{\lambda P\colon}\ |\mathbf{P}| = \mathbf{1}.\ \iota x\ [P(x)]$

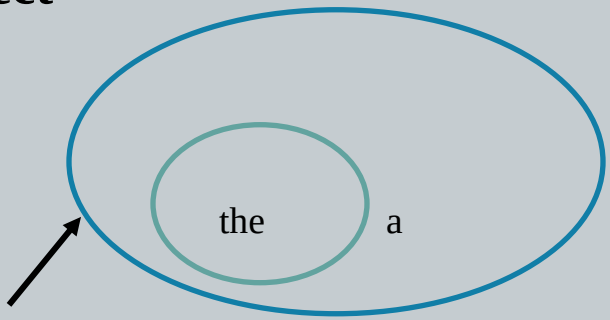
$[\![\text{a}]\!] = \lambda P\ \lambda Q.\ \exists x\ [P(x) \wedge Q(x)]$

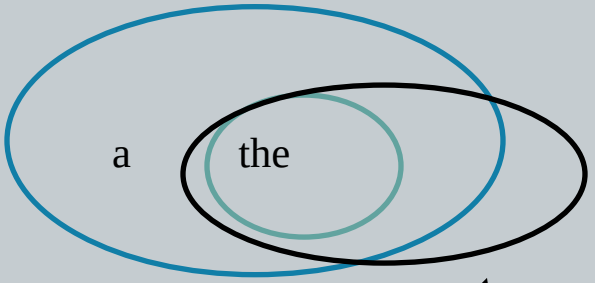
(see also Hawkins 1991)

The Hitchhiker Project

Background Assumptions contd.

The 'standard' view: Bare Nouns in languages without determiners





<u>Dayal 2004 & Dayal 2011</u>
Bare Nouns in languages without determiners are

Like definites in some ways (cf. English *the*),
Like indefinites in some ways (cf. English *a/an*) – but not completely!

They also display properties that neither definites nor indefinites have– namely reference to kinds and differentiated scope.

The Hitchhiker Project

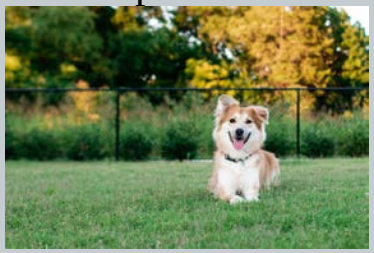
Background Assumptions

Bare NPs in determiner-less languages → a the

Context: Speaker A on phone with Speaker B: Describe what you see outside.

Speaker B: There's **a dog**/***the dog** in the park. ***PIC 1***: *Unique, New info*

Speaker B: There's ??**a dog**/***the dog** in the park. ***PIC 2***: *Non-unique, New info*



Hindi & Russian

10a. ***kutta** maidan meN hai — √maidan meN **kutta** hai — ***PIC 1***: *Unique, New info*
Dog park in is — park in dog is

b. ***kutta** maidan meN hai — *maidan meN **kutta** hai — ***PIC 2***: *Non-unique, New info*
Dog park in is — park in dog is

c. ***Sobaka** sidit na trave — √na trave sidit **sobaka** — ***PIC 1***: *Unique, New info*
Dog sit on grass — on grass sit dog

d. ***Sobaka** sidit na trave — *na trave sidit **sobaka** — ***PIC 2***: *Non-unique, New info*
Dog sit on grass — on grass sit dog — *anita soloveva p.c*

- Hindi and Russian bare singulars encode **uniqueness**, word order signals whether the unique referent is **old information/new information**.

The Hitchhiker Project

The Neo-Carlsonian Approach (Carlson 1989, Chierchia 1998, Dayal 2004):

- Bare nouns are kind terms.
- The same factors regulate the mapping of bare nouns to restrictor/scope that regulate the mapping of indefinites
- Reference to kinds is implicated whether or not the predicate is kind-level.

11a.	Dogs have evolved from wolves.	Evolve-from (∩dogs, ∩wolves)
b.	Dogs bark.	Gen x [∪∩dogs(x)] [bark(x)]
c.	Dogs are barking.	∃x [∪∩dogs(x) ∧ are-barking(x)]

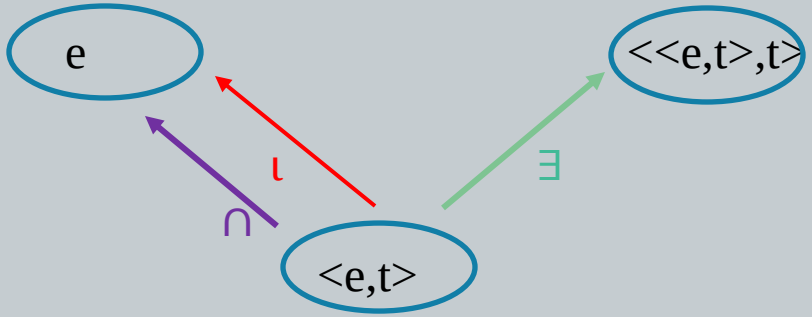
- An alternative view, **the ambiguity view**, takes bare plurals to be kind-denoting with kind-level predicates but property-denoting otherwise:

12a.	Dogs have evolved from wolves.	Evolve-from (∩dogs, ∩wolves)
b.	Dogs bark.	Gen x [dogs(x)] [bark(x)]
c.	Dogs are barking.	∃x [dogs(x) ∧ are-barking(x)]

The Hitchhiker Project

The Neo-Carlsonian Approach: How does a language without determiners navigate the space of kind-reference, definiteness and indefiniteness?

The Partee Triangle (*Partee 1987)*:



Constraints:

Ranking of covertType-shifts: {∩, ι} > ∃

- A bare argument typically will have kind-reference and definite readings, but not indefinite readings -- unless it is the narrowest scope reading based on its profile as a kind term.

Blocking of covert Type-shifts: If lexical D, then no equivalent type-shift.

- A language with a lexical definite determiner will block definite readings: for e.g.English bare plurals typically have kind but not definite readings.

The Hitchhiker Project

The Neo-Carlsonian Approach contd. (Dayal 2004, 2011):

13a. *Dinosaur is extinct. * $\cap(N_{SING})$
b. The dinosaur is extinct. $extinct(\iota x_{TK}(dog^{TK}(x_{TK}))$

- The English singular kind term does not involve *nom.* It is derived from the standard meaning of *the* (iota) applied to a predicate of taxonomic kinds. If iota is covert in a language, singular kind terms will be bare.

Singular kind terms ≠ Plural kind terms.

- They lack ∃ readings in episodic sentences (modulo representative obj readings)
- They may be conceptually plural, but they are not semantically plural:

14a. Dogs are barking outside. $= \exists x[dog(x) \wedge bark(x)]$
b. The dog is barking outside $\neq \exists x[dog(x) \wedge bark(x)]$

15a. The lion gathers near acacia trees. *Krifka et al 1995*
b. The lion attacks each other.

- In languages without determiners, they may allow ***indefinite*** and/or ***number neutral*** readings due to pseudo-incorporation or complex predicate formation – but only in ***direct object position***. (see also *Sag 2019 on Turkish*).

3 Results: A confirmation, a challenge, a discovery

A confirmation: Kind Reference ≈ Differentiated Scope

Bare nouns in all **six** languages showed two properties for the plural N / number neutral base form**: *reference to kinds, differentiated scope*.**

On the neo-Carlsonian approach:

16a. Dogs have evolved from wolves. Evolve-from (∩dogs, ∩wolves)
b. Dogs are barking. ∃x [∪∩dogs(x) ∧ are-barking(x)]
c. Mary baked cakes for three days.
∀t [t ⊆ 3-days → ∃x [∪∩cakes (x) ∧ bake(m,x)

On the ambiguity approach:

17a. Dogs have evolved from wolves. Evolve-from (∩dogs, ∩wolves)
b. Dogs are barking. ∃x [dogs(x) ∧ are-barking(x)]
c. Mary baked cakes for three days.
∀t [t ⊆ 3-days → ∃x [cakes (x) ∧ bake(m,x)

3 Results: A confirmation

On the ambiguity view, where bare plurals (in English) are kind-terms when they are arguments of kind-level predicates, and narrow scope indefinites otherwise, something special needs to be said about differentiated scope readings.

They are **indefinites**, but of a certain very special type –
not like English ***a*** or French ***un/une***
but like French ***du***.

On the neo-Carlsonian view, the differentiated scope reading is derived from the kind-reference inherent to bare plurals. Differentiated scope readings are therefore **predicted** for any language in which the bare argument has the same kind-level profile as English bare plurals.

3 Results: A confirmation, a challenge, a discovery

A challenge: Xhosa bare augmented nouns show the properties associated with **definites & indefinites**.

- They go with globally unique nouns, and are good in bridging contexts.
- They allow wide scope indefinite, narrow scope and differentiated scope readings.

18 a. **I-langa** li-ya khanya namhlanje
AUG-5sun 5SM-DSJ-shine today
"The sun is shining today."

b. U-Helen u-theng-e i-moto. **i-engini** i-nomonakalo
AUG-1Helen 1SM-buy-PST AUG-9car. AUG-9engine 9SM-be_broken
"Helen bought a car. The engine was defective."

c. A-ndi-zange ndi-theng-e **<u>i-n-cwadi</u>**
NEG-1SSA-NEG 1SSM-buy-PST AUG-9-book
Literally: "I didn't buy book." ***∃ > ¬ available***

d. USipho wabulala qho **umvundla**.
AUG-1Sipho 1SM-PST-kill repeatedly AUG-3-rabbit
Literally: "Sipho repeatedly killed rabbit." ***Adv > ∃ available***

These data are from Carstens, Mletshe and Dayal (to appear).

3 Results: A challenge

Does Xhosa show that Ranking of Covert Type-shifts should be abandoned?

Not really!

Two possible explanations:

One, Ranking is parameterized across languages. Russian, Quechua, Korean, Indonesian Cabo Verdean follow ranking, Xhosa doesn't.

A real challenge for Ranking would be a language in which ∃ outranks ∩ and ι.

Two, it is possible that Xhosa does follow ranking but the Xhosa bare augmented noun has a second parse in which ∃ is part of the structure.

This is just a theoretical possibility, I have no empirical evidence to support it.

3 Results: A confirmation, a challenge, a discovery

A discovery: Markers of strict plurality and Sub-types -- *data are from*
Sanchez, Vengoa and Dayal (Quechua),
Choi, Yoon and Dayal (Korean),
Kaufman, Martohardjono and Dayal (Indonesian).

The base form in **Cusco Quechua, Korean** and **Indonesian,** is number-neutral (at least in generic contexts), by the reciprocal test:

19a. Dogs lick each other.
b. *The dog licks each other.

In each language there is a second strategy that leads to a strictly plural interpretation:
-kuna (Quechua), ***-tul*** (Korean), ***reduplication*** (Indonesian).

Does the 2nd strategy involve an operation from #-neutrality to strict plurality?

λP λx [P(x) ∧ ¬AT(x)] **NO!**

3 Results: A discovery

The 2nd strategy resists kind formation & seems to rely on sub-type interpretations:

Quechua-kuna:

20 a. Puma-(kuna)-qa chinka-pu-chka-n-ña-n

puma-PL-TOP lose-BEN-PROG-3.S ADV-FOC/EVID

Without -*kuna*: "Pumas are becoming extinct".

With -*kuna*: "Different types of puma are becoming extinct."

b. dodo (*-kuna) are extinct. *World knowledge: no subtypes of dodos*

Korean -tul:

21a. Konglyong-(tul)-un/i myelcong-toy-ess-ta.

dinosaur-PL-TOP/NOM extinct-become-PST-DECL

Without tul = Dinosaurs are extinct.

With -tul "Different types of dinosaurs became extinct."

b. mammoth/dodo(*-tul) are extinct

World knowledge: no subtypes of mammoth/dodos

3 Results: A discovery

Indonesian reduplication

22a. Dinosaurus (*~dinosaurus) punah
Dinosaur REDUP extinct
"Dinosaurs are extinct."

b. Bambang meng-goreng ikan (~ikan) selama dua hari
Bambang AV-fry fish REDUP for two day
Un-reduplicated: "Bambang fried fish for two days."
Reduplicated: "Bambang fried (different types of) fish for two days"

What is this second strategy? It signals a "***sorting mechanism***". It presupposes that the set it applies to is a plurality whose members can be sorted along the relevant dimension.

Another aspect of such sorters is a tendency to favor ***human*** > ***animate*** > ***inanimate***.

- Tentative idea: this may be related to the fact that humans are perceived as inherently distinct (so sort-able):

23a. A: I met a linguist.
B: Who exactly did you meet?

23b. A: I ate a donut.
B: #What exactly did you eat?

(Dayal & Schwarzschild 2010)

Some Final Take-aways

- The big take-away is that we cannot speak of bare noun phrases as "indefinites" simpliciter. Indefinites come in too many shades for ***bare arguments are indefinites*** to be a meaningful description!

- Claims of ambiguity cannot be based on the bare NP having some properties of *the* and some properties of *a*. **Ambiguous bare arguments** must have ***all* the properties of both**, **not just some of each**. This should lead to their acceptability in every context where either of these lexical options can be used.

- In any discussion of the definite/indefinite status of bare nouns, ***reference to kinds*** must be center-stage, not an add-on.

THANK YOU!